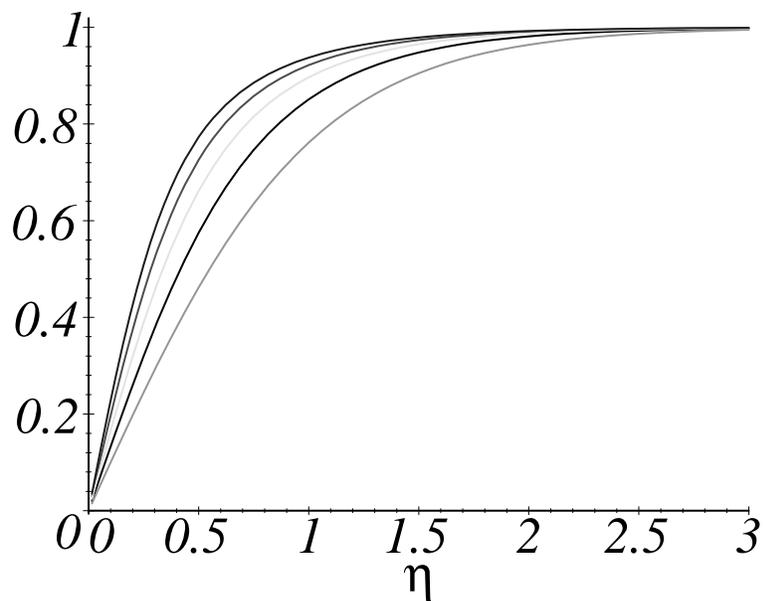# Statistical Physics II (PHYS*4240)

Lecture notes (Fall 2000)

Eric Poisson
Department of Physics
University of Guelph

CONTENTS

CHAPTER 1
REVIEW OF THERMODYNAMICS

1.1   Thermodynamic variables

The purpose of thermodynamics is to describe the properties of various macroscopic systems at, or near, equilibrium. This is done with the help of several *state variables*, such as the internal energy E, the volume V, the pressure P, the number of particles N, the temperature T, the entropy S, the chemical potential μ, and others.

These state variables depend only on *what* the state of the system is, and not on *how* the system was brought to that state. The variables are not all independent, and thermodynamics provides general relations among them. Some of these relations are not specific to particular systems, but are valid for *all* macroscopic systems.

1.2   Equilibrium

A system is in equilibrium if its physical properties do not change with time. The state variables must therefore be constant in time. More precisely, a system is in equilibrium if there are no

- macroscopic motions (mechanical equilibrium)

- macroscopic energy fluxes (thermal equilibrium)

- unbalanced phase transitions or chemical reactions (chemical equilibrium)

occurring within the system.

We will see in Sec. 10 that the requirement for *mechanical* equilibrium is that P must be constant in time and uniform throughout the system. This is easily understood: a pressure gradient produces an unbalanced macroscopic force which, in turn, produces bulk motion within the system; such a situation does not represent equilibrium.

The requirement for *thermal* equilibrium is that T must be constant in time and uniform throughout the system. This also is easily understood: If one part of the system is hotter than the other, energy (in the form of heat) will flow from that part to the other; again such a situation does not represent equilibrium.

Finally, the requirement for *chemical* equilibrium is that μ must be constant in time and uniform throughout the system. The chemical potential is not a very familiar quantity, and this requirement can perhaps not be understood easily. But at least it is clear that unbalanced chemical reactions (we take this to include nuclear reactions as well) cannot occur in a system at equilibrium. Consider for example the reaction of β-decay:

$$n \to p + e^- + \bar{\nu}.$$
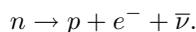
If this reaction is unbalanced (the reversed reaction can occur at sufficiently high densities), the number of neutrons in the system will keep on decreasing, while the number of protons, electrons, and anti-neutrinos will keep on increasing. Such a situation does not represent equilibrium.

A system which is at once in mechanical, thermal, and chemical equilibrium is said to be in *thermodynamic* equilibrium.

1.3 Equation of state

The *equation of state* of a thermodynamic system expresses the fact that not all of its state variables are independent.

Consider a system described by two state variables X and Y, apart from the temperature T. (For example, an ideal gas is described by the variables $X \equiv V$, $Y \equiv P$.) We suppose that the system is initially in equilibrium at temperature T.

We ask the question: Can X and Y be varied *independently*, so that the system remains in equilibrium after the transformation? The answer, which must be determined empirically, is: No! For the system to remain in equilibrium, a given variation of X must be accompanied by a *specific* variation in Y. If Y is not varied by this amount, the system goes out of equilibrium.

There must therefore exist an equation of state,

$$\boxed{f(X, Y, T) = 0}\,, \tag{1.1}$$

relating X, Y, and T for *all* equilibrium configurations. For non-equilibrium states, $f \neq 0$. [For example, the equation of state of an ideal gas is $f(V, P, T) = PV - NkT = 0$.]

The equation of state cannot be *derived* within the framework of thermodynamics. (This is one of the goals of statistical mechanics.) It must be provided as input, and determined empirically.

Suppose the system goes from one equilibrium configuration (X, Y, T) to a neighbouring one $(X + \delta X, Y + \delta Y, T)$. The temperature is kept constant for simplicity; it could also be varied. We can use the equation of state to calculate δY in terms of δX. We have $f(X, Y, T) = 0$ and $f(X + \delta X, Y + \delta Y, T) = 0$, since both these configurations are at equilibrium. So

$$f(X + \delta X, Y + \delta Y, T) - f(X, Y, T) = 0.$$

Expressing the first term as a Taylor series about (X, Y, T), up to first order in the displacements δX and δY, gives

$$\left(\frac{\partial f}{\partial X}\right)_{Y,T} \delta X + \left(\frac{\partial f}{\partial Y}\right)_{X,T} \delta Y = 0.$$

Here, the subscript after each partial derivative indicates which quantities are to be held fixed when differentiating. This equation clearly shows that X and Y cannot be varied independently if the system is to remain in equilibrium during the transformation. (For an ideal gas, the previous equation reduces to $P\delta V + V\delta P = 0$.)

1.4 Quasi-static transformations

In general, changing the state of a thermodynamic system, initially in equilibrium, by a finite amount $(X \to X + \Delta X)$ brings the system out of equilibrium. The free

expansion of a gas, resulting from removing a partition in a partially filled container, is a good example.

However, if the change occurs *sufficiently slowly*, the system will always stay arbitrarily close to equilibrium. The transformation will follow a sequence of equilibrium configurations, along which the equation of state is everywhere satisfied. A good example of such a transformation is the very slow expansion of a gas in a chamber, resulting from the slow displacement of a piston.

Such transformations, which do not take the system away from equilibrium, are called *quasi-static*. We will consider only quasi-static transformations in this course.

1.5   First law of thermodynamics

1.5.1   Formulation

The first law of thermodynamics is a statement of conservation of energy:

> The internal energy E of a system can be increased by *either* letting the system absorb heat (denoted by Q), *or* doing work (denoted by W) on the system.

Mathematically,

$$\boxed{dE = đQ + đW} \,. \tag{1.2}$$

In this equation, $đQ$ and $đW$ are *inexact* differentials. This means that the integrals $\int_A^B đQ$ and $\int_A^B đW$ (where A and B represent the initial and final states of a transformation, respectively) depend not only on the endpoints A and B, but also on the *path* taken to go from A to B. (The path represents the precise way by which the system was taken from state A to state B.) In other words, there are *no* state variables Q and W such that $\int_A^B đQ = Q(B) - Q(A)$ and $\int_A^B đW = W(B) - W(A)$, independently of the path of integration. On the other hand, since E *is* a state variable, the equation $\int_A^B dE = E(B) - E(A)$ is always true, independently of the path.

1.5.2   Work

There are many ways of doing work on a system. One way is to compress (or decompress) the system. A good example of this is the compression of a gas in a chamber, resulting from the very slow (inward) displacement of a piston. The piston moves because of the application of a force F, which is counter-acted by the gas' pressure: $F = -PA$, where A is the piston's area. The work done during a displacement dx is then $đW = F dx = -PA dx$. Since $A dx = dV$, this gives

$$\boxed{đW_{\text{compression}} = -P\, dV} \,. \tag{1.3}$$

The work done on the system is *positive* during a compression; it is *negative* during an expansion.

Another way of doing work is to inject new particles into the system. In this case it must be that $đW \propto dN$. The constant of proportionality defines the *chemical potential* μ, introduced earlier:

$$\boxed{đW_{\text{injection}} = \mu\, dN} \,. \tag{1.4}$$

The chemical potential is just the injection energy per particle.

There are many other forms of work (electric, magnetic, gravitational, shear, etc.). In general, a work term always take the form $f dX$, where X is one of the

physical variables describing the system (such as volume and number of particles), and f is the associated *generalized force* (such as pressure and chemical potential). Work is *always* associated with a change in one of the system's physical variables. (For example, the work done when stretching a rubber band is given by $dW = T d\ell$, where T is the tension and ℓ the length.)

1.5.3   Work depends on path

Consider the quasi-static expansion of an ideal gas, initially in equilibrium. We want to calculate the work done on the system while it is brought from the initial state A to the final state B. Equation (1.3) implies

$$W = -\int_A^B P \, dV.$$

But this cannot be integrated until the function $P(V)$ — the path — is specified. We shall consider two different transformations linking the same initial and final states.

In the first transformation, the gas is expanded at constant temperature T_A. Since the equation of state is always satisfied during a quasi-static transformation, we have $P(V) = NkT_A/V$. Integration yields

$$W_1 = -NkT_A \ln \frac{V_B}{V_A}.$$

In the second transformation, the gas is first expanded at constant pressure P_A, and then its pressure is decreased at constant volume V_B. No work is done during the second part of the transformation. In the first part, $P(V) = P_A = NkT_A/V_A$. Integration yields

$$W_2 = -NkT_A \left(\frac{V_B}{V_A} - 1 \right).$$

We see that the two answers are different, and that the work done on a system does indeed depend on the path.

1.5.4   Heat

By definition, adding heat to a system corresponds to increasing its internal energy *without* doing work, that is, without changing *any* of its physical variables (such as volume, number of particles, etc.). Heat is usually added by putting the system in thermal contact with another system at a higher temperature.

Heat also depends on the path, since work does, but internal energy does not.

1.5.5   Heat capacities

Adding heat to a system usually increases its temperature. A quantitative measure of this is provided by the *heat capacities*, which are defined according to which quantity is held fixed while heat is added.

The *heat capacity at constant volume* is defined as

$$\boxed{C_V = \left(\frac{dQ}{dT} \right)_{V,N}}. \tag{1.5}$$

The *heat capacity at constant pressure* is defined as

$$\boxed{C_P = \left(\frac{dQ}{dT} \right)_{P,N}}. \tag{1.6}$$
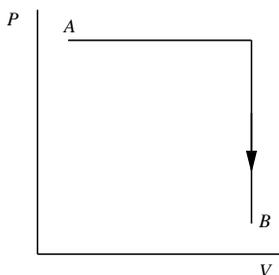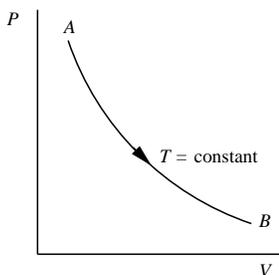
In the framework of thermodynamics, the heat capacities of a given system must be empirically determined (as functions of temperature), and given as input; they cannot be derived. (To theoretically predict the value of the heat capacities is another goal of statistical mechanics.)

However, a general relationship between the heat capacities *can* be derived within thermodynamics, without any need for statistical mechanics. We simply quote the result here:

$$C_P - C_V = -T\left[\left(\frac{\partial V}{\partial T}\right)_P\right]^2\left(\frac{\partial P}{\partial V}\right)_T.$$

The derivation of this result can be found in Sec. 5.7 of Reif's book.

Now, it is true for all substances that $(\partial P/\partial V)_T < 0$: pressure *must* increase as volume is decreased at constant temperature. (The opposite sign would imply instability.) Since the first factor is evidently positive, we conclude that $C_P > C_V$. This is a very general result, which illustrates the power and usefulness of thermodynamics. For an ideal gas, it is easy to use the equation of state $PV = NkT$ to show that $C_P - C_V = Nk$. (Once we have mastered the tools of statistical mechanics, we will be able to calculate that for an ideal gas, $C_P = \frac{5}{2}Nk$ and $C_V = \frac{3}{2}Nk$.)

1.5.6   Heat reservoir

In the sequel we will frequently make use of the idealization of a *reservoir*.

We define a *heat reservoir* (or heat bath) to be a system (in thermal equilibrium) that is so large that to extract any amount of heat from it never affects its temperature. (This works also if heat is delivered to the reservoir.) In other words, a heat reservoir is a system with virtually infinite heat capacity.

We define a *volume reservoir* to be a system (in mechanical equilibrium) that is so large that to vary its volume never affects its pressure.

Finally, we define a *particle reservoir* to be a system (in chemical equilibrium) that is so large that to extract particles from it never affects its chemical potential.

These three idealizations can be combined into a single one, the (unqualified) *reservoir*. This is defined to be a system (in thermodynamic equilibrium) that is so large that interactions with other systems never affect its temperature T, its pressure P, nor its chemical potential μ. It is of course implied that these interactions never take the reservoir away from thermodynamic equilibrium.

1.6   Second law of thermodynamics

1.6.1   Two statements

The second law of thermodynamics is typically formulated in two different ways. The *Kelvin* statement is:

> There exists no thermodynamic transformation whose *sole* effect is to extract a quantity of heat from a system and to convert it entirely into work.

The *Clausius* statement is:

> There exists no thermodynamic transformation whose *sole* effect is to extract a quantity of heat from a colder system and to deliver it to a hotter system.

These two statements have been shown to be completely equivalent.

1.6.2   Reversible, irreversible, and cyclic transformations

A *reversible* transformation from state A to state B is one which can be performed equally well in the opposite direction (from state B to state A), *without introducing any other changes in the thermodynamic system or its surroundings.* As an example of a reversible transformation, consider the quasi-static compression of a gas. The work done on during the compression can be extracted again by letting the gas expand back to its original volume. Since both the gas and the agent responsible for the compression then return to their original states, the transformation is reversible.

An *irreversible* transformation from state A to state B is one which can be performed only in this direction; the reversed transformation would introduce additional changes in the system or its surroundings. As an example of an irreversible transformation, consider the free expansion of a gas caused by the removal of a partition. This transformation is clearly irreversible: replacing the partition once the gas has completely filled the chamber does not make the gas return back to its original volume.

The transfer of heat generally occurs irreversibly. Consider two systems, initially well separated and at different temperatures. The systems are put in thermal contact, and heat flows from the hotter body to the colder one, until they achieve a common temperature. At this point the systems are separated again; this is the final state. The transformation from the initial state to the final state is clearly irreversible: even if the two systems are again put in thermal contact, they have the same temperature, and they can never restore themselves to their initial temperatures. (This would imply a violation of the Clausius statement of the second law.)

In the discussion of the last paragraph, the two systems are completely arbitrary. If one of them turns out to be a heat reservoir, the conclusion would be unaltered. If, however, the reservoir has a temperature that differs by an *infinitesimal* amount from the temperature of the other system, then the transformation would be *reversible*. (We shall prove this in subsection f.) It is therefore possible to add a finite amount of heat to a system in a reversible manner: the system must be put in thermal contact with a succession of heat reservoirs, each at an infinitesimally larger temperature than the previous one. (We show, also in subsection f, that such a procedure is indeed reversible.)

A *cyclic* transformation is one which brings the system back to its original state, irrespective of what may happen to the system's surroundings. In other words, the system's final and initial states are one and the same. Cyclic transformations can be both reversible and irreversible. An example of a reversible cyclic transformation is the quasi-static compression, and then re-expansion, of a gas. An example of an irreversible cyclic transformation is the free expansion of a gas, followed by a slow compression back to its original volume.

1.6.3   Clausius' theorem

This theorem by Clausius follows from the second law:

The inequality

$$\boxed{\oint \frac{dQ}{T} \leq 0}$$ (1.7)

holds for *any* cyclic transformation; equality holds if and only if the transformation is *reversible*.

We will not prove the theorem here, but we will use it to introduce a new state variable, the *entropy*.

1.6.4   Entropy

Consider any cyclic, *reversible* path, taking some thermodynamic system from an initial state A to some state B, and then back to state A. The Clausius theorem implies $0 = \oint đQ/T = \int_A^B đQ/T + \int_B^A đQ/T$, or

$$0 = \int_A^B \frac{đQ}{T} - \int_A^B \frac{đQ}{T}.$$

Here, the two integrals are evaluated along two different paths. If we denote these paths C_1 and C_2, we have

$$\int_{C_1} \frac{đQ}{T} = \int_{C_2} \frac{đQ}{T}.$$

The integral is therefore *independent* of the path taken to go from state A to state B, so long as it is *reversible*.

There must therefore exist a *state variable S* such that $\int_A^B đQ/T = S(B) - S(A)$, independently of the (reversible) path of integration. This state variable is called the *entropy*, and is such that along any reversible path, $dS = đQ/T$ is an exact differential.

The entropy is always defined with respect to some reference state, which we denote O. In terms of this state, the entropy of any other state A is given by

$$\boxed{S(A) = S(O) + \int_R \frac{đQ}{T}}, \tag{1.8}$$

where the path R is *any* reversible path connecting the states O and A.

1.6.5   Properties of the entropy

For *arbitrary* transformations, bringing a thermodynamic system from state A to state B, the change in entropy is related to the integral of $đQ/T$ by

$$\boxed{S(B) - S(A) \geq \int_A^B \frac{đQ}{T}}; \tag{1.9}$$

equality holds for reversible transformations.

The proof of this statement involves Clausius' theorem, with the following cyclic path: the system is first brought from state A to state B along an arbitrary path I, and then brought back to A along a reversible path R. The theorem implies $0 \geq \oint đQ/T = \int_I đQ/T + \int_R đQ/T$. But since $đQ/T = dS$ along R, we have $0 \geq \int_I đQ/T + \int_R dS = \int_I đQ/T + S(A) - S(B)$. This is the same statement as Eq. (1.9).

A special case of this statement holds for an *isolated* system, that is, a system that does not exchange heat with its surroundings. For such a system, Eq. (1.9) implies $S(B) - S(A) \geq 0$, since $đQ \equiv 0$. In other words,

$$\boxed{\text{The entropy of an isolated system } \textit{never} \text{ decreases.}} \tag{1.10}$$

This is often taken as the statement of the second law. As we can see, however, the Kelvin and Clausius statements lead to the more general result (1.9). We note that for reversible transformations involving isolated systems, $S(B) = S(A)$. In other words, *the entropy of an isolated system stays constant during a reversible transformation.*
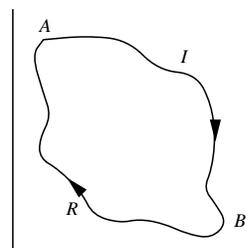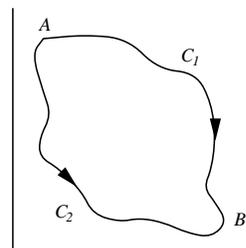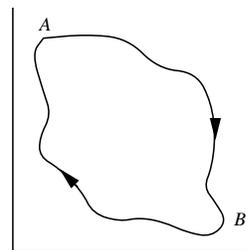
1.6.6   Example: system interacting with a heat reservoir

Consider a system A, which initially is in equilibrium at a temperature T_A, but is then put in thermal contact with a heat reservoir B at a higher temperature T_B. As a result of the interaction, heat flows into A from B, and A's temperature eventually increases to T_B. This transformation is clearly irreversible, and since the combined system $A + B$ is isolated, the change in the total entropy $S_A + S_B$ must be strictly positive.

It is easy to show (this is a homework problem) that for this transformation, the change in total entropy is given by

$$\Delta S = C(x - 1 - \ln x),$$

where C is A's heat capacity, and $x \equiv T_A/T_B$. It is also easy to prove that $x - 1 - \ln x > 0$ for $x < 1$, which establishes that ΔS is strictly positive, as claimed.

Let's consider the special case for which T_B differs from T_A by a very small amount: $T_B = T_A + \delta T$. In this limit, we have $x = 1 - \delta T/T_B$ and $\ln x = -\delta T/T_B - \frac{1}{2}(\delta T/T_B)^2 + \cdots$. This gives

$$\Delta S \simeq \frac{C}{2}\left(\frac{\delta T}{T_B}\right)^2,$$

which is *second order* in the small quantity δT. This proves the statement made earlier, in subsection b, that a system in thermal contact with a heat reservoir at a slightly higher temperature receives heat in a virtually reversible manner.

Any system can be brought *reversibly* from a lower temperature T_1 to a higher temperature T_2. The produce to follow is to first put the system in thermal contact with a slightly hotter reservoir ($T = T_1 + \delta T$), then with another one, slightly hotter than the previous one ($T = T_1 + 2\delta T$), and so on, until the final temperature is reached. At the end of the process, the overall change in total entropy is given by the immediate generalization of the previous result:

$$\Delta S \simeq \frac{C\delta T}{2}\left[\frac{\delta T}{(T_1 + \delta T)^2} + \frac{\delta T}{(T_1 + 2\delta T)^2} + \cdots + \frac{\delta T}{T_2{}^2}\right].$$

The sum within the square brackets is obviously equal to the integral

$$\int_{T_1}^{T_2} \frac{dT}{T^2} = \frac{1}{T_1} - \frac{1}{T_2}$$

when δT is sufficiently small. We finally obtain

$$\Delta S \simeq \frac{C\delta T}{2}\left(\frac{1}{T_1} - \frac{1}{T_2}\right).$$

By choosing δT sufficiently small, ΔS can be made as small as desired. In the limit $\delta T \to 0$, the transformation is reversible.

1.6.7   Third law of thermodynamics

Another property of the entropy, which must derived empirically within the framework of thermodynamics, but can be proven using the tools of statistical mechanics, is known as the third law of thermodynamics. It states:

> The entropy of a thermodynamic system at zero temperature is zero.

We will get back to this later on in the course.

1.7   Thermodynamic potentials

In this and the remaining sections we will consider reversible transformations *only*. For such transformations, it is always true that $đQ = T dS$.

1.7.1   Energy

For reversible transformations, the first law if thermodynamics reads

$$\boxed{dE = T\,dS - P\,dV + \mu\,dN}\,. \tag{1.11}$$

When looked at from a purely mathematical point of view, this equation tells us that E can be expressed as a function of S, V, and N: $E = E(S, V, N)$. Indeed, if such a function exists, then purely mathematically,

$$dE = \left(\frac{\partial E}{\partial S}\right)_{V,N} dS + \left(\frac{\partial E}{\partial V}\right)_{S,N} dV + \left(\frac{\partial E}{\partial N}\right)_{S,V} dN.$$

This has the same form as Eq. (1.11), and it gives us formal *definitions* for T, P, and μ:

$$T \equiv \left(\frac{\partial E}{\partial S}\right)_{V,N}, \qquad P \equiv -\left(\frac{\partial E}{\partial V}\right)_{S,N}, \qquad \mu \equiv \left(\frac{\partial E}{\partial N}\right)_{S,V}.$$

Thus, if E is known as a function of S, V, and N, then the related quantities T, P, and μ can be calculated.

The interpretation of the equation $E = E(S, V, N)$ is that S, V, and N can be considered to be the *independent* state variables, and that E is a *dependent* quantity; T, P, and μ are also dependent quantities. For fixed N, an equilibrium configuration is fully specified by giving the values of S and V. Alternatively, $E = E(S, V, N)$ could be inverted to give $S = S(E, V, N)$, which would make E an independent variable, and S a dependent one. The point is that the four quantities E, S, V, and N are not all independent, and that T, P, and μ can be derived from the equation relating them.

The function $E(S, V, N)$ is the first instance of a *thermodynamic potential*, a function that generates other thermodynamic variables (such as T, P, and μ) by partial differentiation. Other instances are obtained by considering other choices of independent variables. The usefulness of all this will become apparent below.

1.7.2   Enthalpy

We define the enthalpy as

$$\boxed{H = E + PV}\,. \tag{1.12}$$

This equation, when combined with Eq. (1.11), implies

$$dH = T\,dS + V\,dP + \mu\,dN,$$

which states that $H = H(S, P, N)$. We also obtain formal definitions for T, V, and μ:

$$T \equiv \left(\frac{\partial H}{\partial S}\right)_{P,N}, \qquad V \equiv \left(\frac{\partial H}{\partial P}\right)_{S,N}, \qquad \mu \equiv \left(\frac{\partial H}{\partial N}\right)_{S,P}.$$

1.7.3 Helmholtz free energy

We define the Helmholtz free energy as

$$\boxed{F = E - TS}.\qquad(1.13)$$

This equation, when combined with Eq. (1.11), implies

$$dF = -S\,dT - P\,dV + \mu\,dN,$$

which states that $F = F(T, V, N)$. We also obtain formal definitions for S, P, and μ:

$$S \equiv -\left(\frac{\partial F}{\partial T}\right)_{V,N}, \qquad P \equiv -\left(\frac{\partial F}{\partial V}\right)_{T,N}, \qquad \mu \equiv \left(\frac{\partial F}{\partial N}\right)_{T,V}.$$

This choice of independent variables, (T, V, N), seems more natural than (S, V, N) or (S, P, N). We will see in Sec. 11 that the Helmholtz free energy has an important physical meaning; we will encounter it again later.

1.7.4 Gibbs free energy

We define the Gibbs free energy as

$$\boxed{G = E - TS + PV}.\qquad(1.14)$$

This equation, when combined with Eq. (1.11), implies

$$dG = -S\,dT + V\,dP + \mu\,dN,$$

which states that $G = G(T, P, N)$. We also obtain formal definitions for S, V, and μ:

$$S \equiv -\left(\frac{\partial G}{\partial T}\right)_{P,N}, \qquad V \equiv \left(\frac{\partial G}{\partial P}\right)_{T,N}, \qquad \mu \equiv \left(\frac{\partial G}{\partial N}\right)_{T,P}.$$

This choice of independent variables also seems quite natural, and the Gibbs free energy also has an important physical meaning. This will be described in Sec. 11.

1.7.5 Landau potential

As a final example, we consider the Landau potential, which is defined as

$$\boxed{\Omega = E - TS - \mu N}.\qquad(1.15)$$

This equation, when combined with Eq. (1.11), implies

$$d\Omega = -S\,dT - P\,dV - N\,d\mu,$$

which states that $\Omega = \Omega(T, V, \mu)$. We also obtain formal definitions for S, P, and N:

$$S \equiv -\left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu}, \qquad P \equiv -\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu}, \qquad N \equiv -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V}.$$

We will put the Landau potential to good use in a later portion of this course.

1.8   Maxwell relations

The formal definitions for T, S, V, P, μ, and N in terms of the potentials E, H, F, G, and Ω can be used to generate many relations between the various derivatives of these quantities.

For example, Eq. (1.11) implies that $T = \partial E/\partial S$ and $P = -\partial E/\partial V$. This, in turn, gives $\partial T/\partial V = \partial^2 E/\partial V \partial S$. But since the order in which we take the partial derivatives does not matter, we obtain $\partial T/\partial V = \partial^2 E/\partial S \partial V = -\partial P/\partial S$. In other words,

$$\left(\frac{\partial T}{\partial V}\right)_{S,N} = -\left(\frac{\partial P}{\partial S}\right)_{V,N}.$$

This is the first instance of what is known as a Maxwell relation.

Other relations follow just as easily from the other thermodynamic potentials. For example, you may derive these:

$$\left(\frac{\partial T}{\partial P}\right)_{S,N} = \left(\frac{\partial V}{\partial S}\right)_{P,N}, \qquad \left(\frac{\partial S}{\partial V}\right)_{T,\mu} = \left(\frac{\partial P}{\partial T}\right)_{V,\mu},$$

and

$$\left(\frac{\partial S}{\partial P}\right)_{T,N} = -\left(\frac{\partial V}{\partial T}\right)_{P,N}.$$

The Maxwell relations are extremely general, they hold for *any* thermodynamic system. This is another illustration of the power of thermodynamics.

1.9   Scaling properties

Consider two systems in thermodynamic equilibrium. Both systems are identical in every way, except that one is a factor λ larger than the other. We say that the second system is *scaled* by a factor λ with respect to the first system. We ask: How are the state variables of the two systems related? Or, in other words, how do the state variables of a thermodynamic system scale under a scaling of the system?

Clearly, if the system is scaled by a factor λ, it must mean that E, V, and N are all scaled by such a factor: $E \to \lambda E$, $V \to \lambda V$, and $N \to \lambda N$. On the other hand, if the scaled system is otherwise identical to the original system, it must be true that T stays the same: $T \to T$. How about S, P, and μ? The answer comes from the first law: $dE = T\,dS - P\,dV + \mu\,dN$. If E scales but T does not, it must be true that S scales: $S \to \lambda S$. Similarly, if both E and V scale, it must be true that P does not: $P \to P$. And finally, if both E and N scale, it must be true that μ does not: $\mu \to \mu$.

We can therefore separate the thermodynamic variables into two groups. The first group contains the variables that scale under a scaling of the system; such variables are called *extensive*. The second group contains the variables that do not scale under a scaling of the system; such variables are called *intensive*. We have found that E, S, V, and N are extensive variables, and that T, P, and μ are intensive variables.

These scaling properties lead to an interesting consequence when we recall that $E = E(S, V, N)$. Upon scaling, this equation implies

$$E(\lambda S, \lambda V, \lambda N) = \lambda E(S, V, N);$$

whatever the function $E(S, V, N)$ happens to be, it must satisfy this scaling equation. Let $\lambda = 1+\epsilon$, where $\epsilon \ll 1$. Then $E(S+\epsilon S, V+\epsilon V, N+\epsilon N) = (1+\epsilon)E(S, V, N)$.

Expanding the left-hand side of this equation in powers of ϵ, up to first order, we arrive at

$$\left(\frac{\partial E}{\partial S}\right)_{V,N} S + \left(\frac{\partial E}{\partial V}\right)_{S,N} V + \left(\frac{\partial E}{\partial N}\right)_{V,N} N = E.$$

Or, recalling the formal definitions for T, P, and μ,

$$\boxed{E = TS - PV + \mu N} \ . \tag{1.16}$$

Now for another interesting result. If we differentiate Eq. (1.16) and substitute Eq. (1.11), we quickly obtain

$$\boxed{N\,d\mu = -S\,dT + V\,dP} \ . \tag{1.17}$$

This states that the intensive variables are not all independent: $\mu = \mu(T, P)$. Equation (1.17) is known as the *Gibbs-Duhen relation*.

1.10   Equilibrium conditions for isolated, composite systems

We now turn from formal considerations to more physical ones. In this section we consider two thermodynamic systems, A and B, which initially are separated and in equilibrium. The systems are then brought together to form a composite system C, and are allowed to interact. In general, C is not initially in equilibrium. For example, if A and B are at a different temperature, heat will flow from the hotter part of C to the colder part, and this cannot represent an equilibrium situation. However, interactions between the two subsystems will give rise to an irreversible transformation from C's original state to a final, equilibrium state. In our example, A and B will eventually reach a common temperature, at which point equilibrium will be re-established.

We wish to derive the conditions under which a composite system will be in equilibrium. We proceed as follows:

Because C is an isolated system and the transformation is irreversible, its entropy must be increasing during the transformation:

$$dS_C = dS_A + dS_B \geq 0;$$

the equality sign takes over when equilibrium is established. Now, the change in entropy in subsystem A, dS_A, can be computed by considering *any* reversible transformation that relates the same initial and final states as the *actual*, irreversible transformation: we just apply Eq. (1.8). Along such a path, $T_A\,dS_A = đQ_A = dE_A - đW_A = dE_A + P_A\,dV_A - \mu_A\,dN_A$. Similarly, $T_B\,dS_B = dE_B + P_B\,dV_B - \mu_B\,dN_B$. The final step of the derivation is to notice that since C is an isolated system, its total energy must be conserved: $0 = dE_C = dE_A + dE_B$. Similarly, $dV_A + dV_B = 0$, and $dN_A + dN_B = 0$. Combining these results, we easily obtain the fundamental equation

$$\boxed{\left(\frac{1}{T_A} - \frac{1}{T_B}\right) dE_A + \left(\frac{P_A}{T_A} - \frac{P_B}{T_B}\right) dV_A - \left(\frac{\mu_A}{T_A} - \frac{\mu_B}{T_B}\right) dN_A \geq 0} \ . \tag{1.18}$$

As before, equality holds at equilibrium.

To extract the equilibrium conditions from Eq. (1.18), we consider some special cases.

We first consider a purely *thermal* interaction between the subsystems A and B. This means that the partition separating A and B allows for the exchange

of heat, but it cannot move, and it does not allow particles to pass through. In other words, the interaction leaves V_A and N_A fixed. Equation (1.18) then implies $(1/T_A - 1/T_B)dE_A \geq 0$. During the approach to equilibrium, the inequality sign applies. We therefore find that $dE_A > 0$ if $T_A < T_B$, while $dE_A < 0$ if $T_A > T_B$. This certainly conforms to our physical intuition. On the other hand,

$$T_A = T_B$$

at equilibrium. This also conforms to physical intuition.

We next consider a purely *mechanical* interaction between A and B, now imagined to have equal temperatures. This means that the partition separating A and B is now able to move, so that V_A might change during the interaction. In other words, the interaction leaves only N_A fixed, and $T_A = T_B$. For such a situation, Eq. (1.18) implies $(P_A - P_B)dV_A \geq 0$. The inequality sign holds during the approach to equilibrium, and we find that $dV_A > 0$ if $P_A > P_B$, while $dV_A < 0$ if $P_A < P_B$. This makes sense: the subsystem with the larger pressure pushes the partition outward. We also find that

$$P_A = P_B$$

at equilibrium, which also makes good physical sense.

Finally, we consider a purely *chemical* interaction between A and B, now imagined to have identical temperatures and pressures ($T_A = T_B$ and $P_A = P_B$). This means that particles can now pass through the partition separating A from B. Equation (1.18) now gives $(\mu_A - \mu_B)dN_A \leq 0$. This means that a difference in the chemical potentials gives rise to a flux of particles from one subsystem to the other. This also implies that

$$\mu_A = \mu_B$$

at equilibrium.

To summarize, we have found that the composite system achieves equilibrium when, as a result of the interaction between the subsystems, the temperatures, pressures, and chemical potentials have all equalized. This means that at equilibrium, the temperature, pressure, and chemical potential must be *uniform* throughout the system. This conclusion holds for any composite system in isolation; the number of parts is irrelevant. This conclusion also holds for any isolated system, since such a system can always be partitioned into two or more parts (this partitioning can be real or fictitious). We therefore have derived the equilibrium conditions first introduced in Sec. 2.

1.11   Equilibrium for interacting systems

The equilibrium conditions derived in the preceding section are appropriate for isolated systems. Here we wish to derive equilibrium conditions that are appropriate for systems interacting with a *reservoir*. We recall from Sec. 5f that a reservoir is a system so large that interactions with other systems never affect its *intensive* variables (temperature, pressure, and chemical potential).

We consider a system A, initially isolated and in equilibrium, but then made to interact with a reservoir R at temperature T_R, pressure P_R, and chemical potential μ_R. The combined system $C = A + R$, which is isolated, is initially not in equilibrium. However, interactions between A and R give rise to an irreversible transformation from C's original state to its final, equilibrium state. We seek the conditions that hold when equilibrium is established.

The derivation proceeds much as before. We first use the fact that the combined system is isolated to write $dE_A + dE_R = dV_A + dV_R = dN_A + dN_R = 0$. Next, we

calculate that the change in the reservoir's entropy is given by

$$
\begin{aligned}
T_R\,dS_R &= dE_R + P_R\,dV_R - \mu_R\,dN_R \\
&= -dE_A - P_R\,dV_A + \mu_R\,dN_A.
\end{aligned}
$$

Next, we write $T_R(dS_A + dS_R) = T_R\,dS_A - dE_A - P_R\,dV_A + \mu_R\,dN_A$, and notice that the left-hand side must be non-negative. This gives us our fundamental equation:

$$
\boxed{dE_A - T_R\,dS_A + P_R\,dV_A - \mu_R\,dN_A \le 0}\,. \tag{1.19}
$$

This equation holds for *any* kind of interaction between a system A and a reservoir R; equality holds at equilibrium.

We now consider a few special cases.

First, we consider an *isolated system*. An isolated system can always be considered to be interacting, though the "interaction" involves no change in energy, volume, or number of particles. In other words, we consider an "interaction" that keeps E_A, V_A, and N_A fixed. For such a case Eq. (1.19) implies $dS_A \ge 0$, which is nothing but the compact mathematical statement of the second law of thermodynamics. In other words:

> For interactions that leave E, V, and N fixed — that is, for isolated systems — the entropy S is maximized at equilibrium.

Second, we consider a purely *thermal* interaction, which keeps V_A and N_A fixed. For such a case Eq. (1.19) implies $d(E_A - T_R S_A) \le 0$, which states that at equilibrium, the quantity $E_A - T_R S_A$ is minimized. But since we already know that temperature must be uniform at equilibrium ($T_A = T_R$), we obtain the following statement:

> For interactions that leave V and N fixed — that is, for thermal interactions — the Helmholtz free energy $F = E - TS$ is minimized at equilibrium.

Third, we consider a *thermo-mechanical* interaction, during which only N_A is kept fixed. For such a case Eq. (1.19) implies $d(E_A - T_R S_A + P_R V_A) \le 0$, which states that at equilibrium, the quantity $E_A - T_R S_A + P_R V_A$ is minimized. But since we know that temperature and pressure must both be uniform at equilibrium ($T_A = T_R$, $P_A = P_R$), we obtain the following statement:

> For interactions that leave N fixed — that is, for thermo-mechanical interactions — the Gibbs free energy $G = E - TS + PV$ is minimized at equilibrium.

Fourth, and last, we consider a *thermo-chemical* interaction, during which only V_A is kept fixed. For this case Eq. (1.19) implies $d(E_A - T_R S_A - \mu_R N_A) \le 0$, which states that at equilibrium, the quantity $E_A - T_R S_A - \mu_R N_A$ is minimized. But since we know that temperature and chemical potential must both be uniform at equilibrium ($T_A = T_R$, $\mu_A = \mu_R$), we obtain the following statement:

> For interactions that leave V fixed — that is, for thermo-chemical interactions — the Landau potential $\Omega = E - TS - \mu N$ is minimized at equilibrium.

These various equilibrium conditions give physical meaning to the thermodynamic potentials introduced in Sec. 7.

1.12   Limitations of thermodynamics

The laws of thermodynamics can be used to derive very general relations among the various state variables and their derivatives. These relations are not specific to particular systems, but hold for *all* macroscopic systems at equilibrium.

However, because thermodynamics makes no contact with the microphysics of the systems it studies, it cannot give a complete picture. For example, thermodynamics alone cannot provide an expression for a system's equation of state, nor can it provide the function $S(E, V, N)$, and it cannot provide values for the heat capacities.

The role of statistical mechanics is to make contact with the microphysics. Its role is not to *derive* the laws of thermodynamics. In fact, thermodynamics stands very well on its own. The role of statistical mechanics is to *extend* and *complement* the framework of thermodynamics. It also makes the language of thermodynamics much more concrete, by providing a way of *computing* the equation of state, the function $S(E, V, N)$, and the heat capacities.

1.13   Brief summary of thermodynamics

- The *first law of thermodynamics* states that the internal energy E of a system can be increased by either adding heat or doing work. Mathematically,

$$dE = đQ + đW.$$

  Here, $đQ$ and $đW$ are *inexact differentials*, meaning that the integrals $\int_A^B đQ$ and $\int_A^B đW$ depend not only on the endpoints, but also on the path taken to go from state A to state B. On the other hand, because E is a *state function*, $\int_A^B dE = E(B) - E(A)$, independently of the path.

- The *second law of thermodynamics* states that there exists a state function S, called the entropy, which satisfies the inequality

$$dS \geq \frac{đQ}{T};$$

  equality holds for *reversible* transformations. For *thermally isolated* systems (such that $đQ \equiv 0$), the entropy satisfies $dS \geq 0$: The entropy of an isolated system can never decrease, and stays constant during reversible transformations.

- When heat is applied to a system, its temperature generally increases. A quantitative measure of this is given by the *heat capacities*:

$$C_V = \left(\frac{đQ}{dT}\right)_{V,N}, \qquad C_P = \left(\frac{đQ}{dT}\right)_{P,N}.$$

  It is true *for all* thermodynamic systems that $C_P > C_V$.

- During reversible transformations ($đQ = TdS$), the first law of thermodynamics reads

$$dE = T\,dS - P\,dV + \mu\,dN.$$

  This shows that E can be regarded as a function of S, V, and N. If the function $E(S, V, N)$ is known, then it can be used to *define* T, P, and μ:

$$T = \left(\frac{\partial E}{\partial S}\right)_{V,N}, \quad P = -\left(\frac{\partial E}{\partial V}\right)_{S,N}, \quad \mu = \left(\frac{\partial E}{\partial N}\right)_{S,V}.$$

The internal energy is one of the *thermodynamic potentials*. The others are the enthalpy H, the Helmhotz free energy F, the Gibbs free energy G, and the Landau potential Ω. They are defined by

$$
\begin{aligned}
H &= E + PV & dH &= TdS + VdP + \mu dN \\
F &= E - TS & dF &= -SdT - PdV + \mu dN \\
G &= E - TS + PV & dG &= -SdT + VdP + \mu dN \\
\Omega &= E - TS - \mu N & d\Omega &= -SdT - PdV - Nd\mu
\end{aligned}
$$

The thermodynamic potentials can be used to derive the *Maxwell relations*.

- Thermodynamic variables can be divided into two groups, depending on their behaviour under a *scaling* of the system. *Extensive* variables scale with the system; E, S, V, and N are extensive variables. *Intensive* variables stay invariant; T, P, and μ are intensive variables.

- An isolated system is in *thermal* equilibrium (no macroscopic energy fluxes) if T is uniform throughout the system. It is in *mechanical* equilibrium (no macroscopic motions) if P is uniform. And it is *chemical* equilibrium (no unbalanced phase transitions or chemical reactions) if μ is uniform. If a system is at once in thermal, mechanical, and chemical equilibrium, then it is said to be in *thermodynamic* equilibrium.

- A *reservoir* is a system in thermodynamic equilibrium that is so large that interactions with other systems *never* produce any change in its intensive variables T, P, and μ. Its extensive variables, however, may change during the interaction.

- During the interaction of a system [with variables (E, T, S, P, V, μ, N)] with a reservoir [with variables (T_0, P_0, μ_0)], the inequality

$$
dE - T_0 dS + P_0 dV - \mu_0 dN \leq 0
$$

  always hold; equality holds when equilibrium is established. This inequality can be used to prove that: (i) for isolated systems, equilibrium is achieved when S reaches a *maximum*; (ii) for interactions that keep V and N fixed, equilibrium is achieved when $T = T_0$ and $F = E - TS$ reaches a *minimum*; (iii) for interactions that keep N fixed, equilibrium is achieved when $T = T_0$, $P = P_0$, and $G = E - TS + PV$ reaches a *minimum*; and (iv) for interactions that keep V fixed, equilibrium is achieved when $T = T_0$, $\mu = \mu_0$, and $\Omega = E - TS - \mu N$ reaches a *minimum*.

1.14   Problems

1. (Adapted from Reif 2.5 and 2.6) The purpose of this problem is to look at the differences between exact and inexact differentials from a purely mathematical point of view.

   a) Consider the infinitesimal quantity $Adx + Bdy$, where A and B are functions of x and y. If this quantity is to be an *exact* differential, then there must exist a function $F(x, y)$ whose total derivative is given by $dF = Adx + Bdy$. Show that in this case, A and B must satisfy the condition

$$
\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}.
$$

   If, on the other hand, $Adx + Bdy$ is an *inexact* differential, then there exists no such function $F(x, y)$, and this condition need not hold. So this

equation provides a useful way of testing whether or not some infinitesimal quantity is an exact differential.

b) Let $đF = (x^2 - y)dx + xdy$. Is $đF$ an exact differential?

c) Let $đG = đF/x^2$, with $đF$ as given in part b). Is $đG$ an exact differential?

d) Prove that if $Adx + Bdy$ is an exact differential, in the sense of part a), then its integral from (x_1, y_1) to (x_2, y_2) depends on the endpoints only, and not on the path taken to go from one point to the other.

2. (Reif 2.10) The pressure of a thermally insulated amount of gas varies with its volume according to the relation $PV^\gamma = K$, where γ and K are constants. Find the work done *on* this gas during a quasi-static process from an initial state (P_i, V_i) to a final state (P_f, V_f). Express your result in terms of P_i, P_f, V_i, V_f, and γ.

3. (Reif 2.11) In a quasi-static process $A \to B$ (see diagram) during which no heat is exchanged with the environment (this is called an *adiabatic process*), the pressure of a certain amount of gas is found to change with its volume V according to the relation $P = \alpha V^{-5/3}$, where α is a constant.

Find the work done *on* the system, and the net amount of heat absorbed by the system, in each of the following three processes, all of which takes the system from state A to state B. Express the results in terms of P_A, P_B, V_A, and V_B.

a) The system is expanded from its original volume to its final volume, heat being added to maintain the pressure constant. The volume is then kept constant, and heat is extracted to reduce the pressure to its final value.

b) The volume is increased and heat is supplied to cause the pressure to decrease linearly with the volume.

c) The steps of process a) are carried out in the opposite order.

4. (Reif 4.1) To do this problem you need to know that the heat capacity C of a kilogram of water is equal to 4.18×10^3 J/K. The heat capacity is assumed not to vary with temperature.

a) One kilogram of water at $0°C$ is brought into contact with a large heat reservoir at $100°C$. When the water has reached $100°C$, what has been the change in entropy of the water? Of the reservoir? Of the entire system consisting of both the water and the reservoir?

b) If the water had been heated from $0°C$ to $100°C$ by first bringing it in contact with a reservoir at $50°C$ and then with a reservoir at $100°C$, what would have been the change in entropy of the entire system?

c) Show how the water might be heated from $0°C$ to $100°C$ with no change in the entropy of the entire system.

5. (Reif 4.3) The heat absorbed by a certain amount of ideal gas during a quasi-static process in which its temperature T changes by dT and its volume V by dV is given by $đQ = CdT + PdV$, where C is its constant heat capacity at constant volume, and $P = NkT/V$ is the pressure.

Find an expression for the change in entropy of this gas, during a process which takes it from initial values of temperature T_i and volume V_i to final values T_f and V_f. Does the answer depend on the kind of process involved in going from the initial state to the final state? (Explain.)
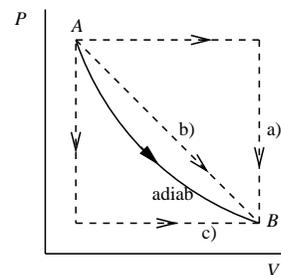
6. A thermodynamic system A contains n times as many particles as system B, so that their heat capacities (at constant volume) are related by $C_A = nC_B$. Initially, both systems are isolated from each other, and at temperatures T_A and T_B ($T_A > T_B$), respectively. The systems are then brought into thermal contact; their respective volumes do not change during the interaction. After equilibrium is re-established, the systems are separated again, and are found to be at a common temperature T_F.

   a) Calculate T_F.

   b) Calculate ΔS, the amount by which the total entropy has increased during the interaction.

   c) Consider the case $n \gg 1$. Show that in this case, your expressions for T_F and ΔS reduce to

$$
\begin{aligned}
T_F &\simeq T_A\left[1 - \frac{1}{n}(1 - x)\right], \\
\Delta S &\simeq C_B(x - 1 - \ln x),
\end{aligned}
$$

   where $x = T_B/T_A$. [Hint: Use the approximation $\ln(1 + \epsilon) \simeq \epsilon$, valid for $\epsilon \ll 1$.]

CHAPTER 2
STATISTICAL MECHANICS OF
ISOLATED SYSTEMS

2.1 Review of probabilities

Not surprisingly, statistical mechanics makes rather heavy use of probability theory. Fortunately, however, only the most primitive concepts are required. We review these concepts here.

2.1.1 Probabilities

Consider a *random process* that returns *any* one of the following numbers: $\{x_1, x_2, \cdots, x_M\}$. We assume that the numbers are ordered ($x_1 < x_2 < \cdots < x_M$). We also assume that the process can be repeated any number of times, and that each result is independent of the previous one.

After N applications of the process, it is found that the particular number x_r is returned N_r times. Clearly, $N = N_1 + N_2 + \cdots + N_m$. The *probability* that the number x_r will be returned is defined by

$$p_r = \lim_{N \to \infty} \frac{N_r}{N}, \qquad \sum_{r=1}^{M} p_r = 1.$$

Probabilities obey the following basic rules:

The "and" rule: After two applications of the random process, the probability that *both x_r and $x_{r'}$* are returned is equal to $p_r \times p_{r'}$. This generalizes to an arbitrary number of applications:

$$P(x_r \text{ and } x_{r'} \text{ and } \cdots) = p_r \times p_{r'} \times \cdots.$$

The "or" rule: After a single application of the random process, the probability that *either x_r or $x_{r'}$* are returned is equal to $p_r + p_{r'}$. This generalizes to

$$P(x_r \text{ or } x_{r'} \text{ or } \cdots) = p_r + p_{r'} + \cdots.$$

2.1.2 Averages

After N applications of the random process, the *average value* of the random variable is given by $\overline{x} = (x_1 N_1 + x_2 N_2 + \cdots + x_M N_M)/N$. In the limit $N \to \infty$, we obtain

$$\overline{x} = \sum_{r=1}^{M} x_r \, p_r.$$

19

This result generalizes to any function $f(x_r)$ of the random variable:

$$\overline{f(x)} = \sum_{r=1}^{M} f(x_r) p_r.$$

In particular, $\overline{x^2} = \sum_{r=1}^{M} x_r^2 p_r$. Notice that $\overline{x^2} \neq \overline{x}^2$. The difference between these two quantities is the square of the *standard deviation* σ:

$$\sigma^2 = \overline{(x - \overline{x})^2} = \sum_{r=1}^{M} (x_r - \overline{x})^2 p_r.$$

It is indeed easy to show that $\sigma^2 = \overline{x^2} - \overline{x}^2$.

2.1.3 Continuous variables

Up to this point we have considered a random variable that takes only a discrete set of values. We now generalize to a *continuous* random variable. As a guide for taking the continuum limit, consider, for a *discrete* variable,

$$
\begin{aligned}
P(x_{r'} \leq x \leq x_{r''}) \;&=\; \text{probability that the random variable is found} \\
&\qquad \text{in the interval between } x_{r'} \text{ and } x_{r''} \\
&=\; P(x_{r'} \text{ or } x_{r'+1} \text{ or } \cdots \text{ or } x_{r''}) \\
&=\; p_{r'} + p_{r'+1} + \cdots + p_{r''} \\
&=\; \sum_{r=r'}^{r''} p_r.
\end{aligned}
$$

In the continuous limit, we would like to replace this by an integral between the two values $x_{r'}$ and $x_{r''}$. However, the sum appearing above is a sum over the index r, and not a sum over the variable itself. It is easy to remedy this:

$$\sum_{r=r'}^{r''} p_r = \sum_{x=x_{r'}}^{x_{r''}} \frac{p_r}{\Delta x_r} \Delta x_r,$$

where $\Delta x_r \equiv x_{r+1} - x_r$ is the increment in x when r is increased by one unit. (Δx_r may be different for each value of r; the values of the random variable are not necessarily equally spaced.)

Passage to the continuous limit is now obvious. Setting $x_1 \equiv x_{r'}$, $x_2 \equiv x_{r''}$, and defining the function $p(x)$ as the continuous limit of $p_r / \Delta x_r$, we arrive at

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} p(x) dx.$$

The function $p(x)$ is called the *probability distribution function* of the random variable x. Its meaning is obtained by considering a special case of the previous equation: $P(x_0 \leq x \leq x_0 + \delta x_0) = \int_{x_0}^{x_0 + \delta x_0} p(x)\, dx$, where δx_0 is an infinitesimal quantity. The integral is obviously equal to $p(x_0)\delta x_0$. Changing the notation slightly, this gives us the following interpretation:

$$
\begin{aligned}
p(x)dx \;&=\; \text{probability that the random variable is found} \\
&\qquad \text{in the interval between } x \text{ and } x + dx.
\end{aligned}
$$

Probabilities are normalized also for continuous variables:

$$\int p(x)dx = 1.$$

Here, the integral extends over all possible values of the random variable. Averages are also computed in the obvious way:

$$\overline{f(x)} = \int f(x)\,p(x)dx.$$

This concludes our review of probability theory.

2.2   Macrostate and microstates

We now begin our formulation of statistical mechanics for systems in isolation. The first order of business is to introduce two very different ways of looking at a thermodynamic system. The first way is *macroscopic*, and employs classical ideas. The second way is *microscopic*, and employs quantum ideas.

The macroscopic description of a system is made by specifying what we shall call its *macrostate*. The macrostate of an isolated system is specified by giving the values of just three quantities: E, V, and N. As we saw in part A of this course, all other thermodynamic variables can be determined, at least in principle, from these three fundamental quantities.

Actually, it is usually more appropriate (and also mathematically more convenient) to specify the system's energy within a narrow range, $E < \text{energy} < E + \delta E$, where $\delta E \ll E$. The quantity δE represents the unavoidable uncertainty in measuring the energy of the system. This uncertainty can be associated with experimental limitations, or with the Heisenberg uncertainty principle.

The microscopic description of a system is made by specifying what we shall call its *microstate*. The microstate of an isolated system is specified by giving the quantum mechanical wave function which describes the system. The wave function depends on all of the system's generalized coordinates; we write this dependence as $\psi = \psi(q_1, q_2, \ldots, q_f)$, where f denotes the system's total number of degrees of freedom. Equivalently, the microstate can be specified by giving the values of all of the f quantum numbers which characterize the system's quantum state. We will label the possible microstates of a system by the abstract index r; in effect,

$$r = \{f \text{ quantum numbers}\}.$$

(In words, this reads "r is the set of all f quantum numbers".) The corresponding wave function will be denoted ψ_r.

It should be clear that the microstate gives a very detailed description of the system. A typical macroscopic system possesses a number N of particles of the order of 10^{23}. Since each particle possesses at least three degrees of freedom (and more if the particle has spin), the total number of quantum numbers that must be specified in order to fully characterize this system's microstate is at least $3N$, which also is of order 10^{23}. To give such a detailed description of a macroscopic system is clearly impractical, and this is why we must resort to statistical methods. On the other hand, the system's macrostate is fully determined once three quantities (E, V, and N) are specified. This is a very coarse description, but one which is much more practical to give.

A simple example of a quantum mechanical system is the simple harmonic oscillator (SHO). The quantum states of this system are fully characterized by a single quantum number n, which takes nonnegative integer values ($n = 0$ corresponds to the ground state). The energy eigenvalues are $E_n = (n + \frac{1}{2})\hbar\omega$, where ω is the characteristic frequency of the oscillator. The corresponding eigenfunctions can be written in terms of Hermite polynomials.

A more complicated example of a quantum system is one which contains N identical SHOs. The quantum state is now specified by giving the value of each one

of the N quantum numbers. We write $r = \{n_1, n_2, \ldots, n_N\}$, the set of all individual quantum numbers. The energy of such a state r is

$$E_r = (n_1 + \tfrac{1}{2})\hbar\omega + (n_1 + \tfrac{1}{2})\hbar\omega + \cdots + (n_N + \tfrac{1}{2})\hbar\omega,$$

and the wave function for the entire system is obtained by taking the product of all individual wave functions.

Another example of a quantum system is that of a free particle confined to a box of volume V. Here, three quantum numbers are necessary to specify the system's quantum state; these take positive integer values. We thus write $r = \{n_x, n_y, n_z\}$, and

$$E_r = \frac{(2\pi\hbar)^2}{8mV^{2/3}} \left(n_x{}^2 + n_y{}^2 + n_z{}^2\right)$$

are the energy eigenvalues (m is the mass of the particle).

The generalization to N identical (noninteracting) particles is obvious. Here, $3N$ quantum numbers must be specified in order to fully characterize the system's microstate. We therefore write $r = \{n_i\}$, the set of all quantum numbers n_i, where $i = 1, 2, \ldots, 3N$ labels all of the system's $3N$ degrees of freedom. The energy eigenvalues are

$$\boxed{E_r = \frac{(2\pi\hbar)^2}{8mV^{2/3}} \sum_{i=1}^{3N} n_i{}^2} \,. \tag{2.1}$$

This quantum system will form our physical model for an ideal gas.

2.3   Statistical weight

2.3.1   Definition

An isolated system in a specified macrostate (E, V, N) can be found to be in *many* different microstates. As a very simple example, consider a system consisting of 2 SHOs whose total energy is known to be equal to $10\hbar\omega$ (hardly a macroscopic value!). We want to find the number of microstates accessible to this system. Clearly, the sum $n_1 + n_2$ must then be equal to 9, and it is easy to see that there are 10 different ways of choosing n_1 and n_2 such that $n_1 + n_2 = 9$. So the number of accessible microstates is equal to 10.

The number of microstates accessible to a thermodynamic system in a macrostate (E, V, N) is called the *statistical weight* of this system, and is denoted $\Omega(E, V, N)$. We shall see that calculation of the statistical weight gives access to all of the thermodynamic properties of a system. Strictly speaking, the statistical weight depends also on δE, the energy uncertainty; a better notation for it would therefore be $\Omega(E, V, N; \delta E)$. We shall see that typically (but not always!), the statistical weight is *proportional* to δE: When $\delta E \ll E$,

$$\boxed{\Omega(E, V, N; \delta E) = \omega(E, V, N)\, \delta E} \,. \tag{2.2}$$

The quantity $\omega(E, V, N)$ is called the *density of states*, and it is independent of δE.

2.3.2   Example: single particle in a box

Before we can relate the statistical weight to thermodynamics, we must first learn how to compute it. We first do so for a system consisting of a single particle of mass m in a box of volume V. The system will be in the macrostate $(E, V, 1)$ if its

energy eigenvalue satisfies the condition $E < E_r < E + \delta E$. (An expression for the eigenvalues was worked out in Sec. 2.) If we define the quantities

$$R = \frac{2\sqrt{2mE}\,V^{1/3}}{2\pi\hbar}, \qquad \delta R = \frac{R\,\delta E}{2E},$$

then a little algebra shows that this condition can be written as

$$R < \sqrt{n_x{}^2 + n_y{}^2 + n_z{}^2} < R + \delta R.$$

This has a nice geometric interpretation. If we imagine a fictitious three-dimensional space spanned by the directions n_x, n_y, and n_z, then the square root represents a distance in this space. The inequality therefore states that this distance must be bounded by two values, R and $R + \delta R$. Geometrically, this condition describes a *spherical shell* of radius R and thickness δR. To compute Ω amounts to counting the number of points within the shell.

The situation is made complicated by two different factors. First, we must recall that n_x, n_y, and n_z take positive values only. So our fictitious space is not quite like ordinary 3D space, which is infinite in all directions, positive and negative. Instead, it is *semi-infinite*, that is, infinite only in the positive directions. We are therefore dealing with the *first octant* of ordinary 3D space. The geometric figure described by the inequality is therefore not that of a complete spherical shell, but only *one octant* thereof. The second complication is that our fictitious space is not a continuum, like ordinary 3D space, but a discrete space. This makes counting the points within the spherical shell a bit difficult. To remove this difficulty, we will assume that the system's energy E is truly macroscopic, so that the quantum numbers n_x, n_y, and n_z must all be very large. In this limit, the discrete nature of the space becomes irrelevant: as we move from one point to the next, the increment in n_x (say) is much smaller than n_x itself, $\Delta n_x/n_x \ll 1$. In this limit, therefore, we can treat our discrete space as one octant of ordinary three-dimensional space, and to count the number of points within the shell amounts to computing its *volume*.

This is easy. The volume of a three-dimensional sphere of radius R is $V_3(R) = \frac{4}{3}\pi R^3$, and the volume of a shell of radius R and thickness δR must then be $\delta V_3 = V_3(R + \delta R) - V_3(R) = (dV_3/dR)\delta R = 4\pi R^2 \delta R$. We must still account for the fact that only one octant of the shell is actually there; this we do by simply dividing the result by $2^3 = 8$, a factor of two for each dimension. Finally, we arrive at

$$\Omega(E, V, 1; \delta E) = \frac{1}{2^3}\,4\pi R^2 \delta R = \frac{2\pi V}{(2\pi\hbar)^3}\,(2mE)^{3/2}\,\frac{\delta E}{E}.$$

We see that the statistical weight is indeed of the form (2.2), with the density of states given by everything that multiplies δE.

2.3.3   Example: N particles in a box

It is not too hard to generalize the previous discussion to the case of N noninteracting particles. The energy eigenvalues for this system are given by Eq. (2.1). It is easy to show that in terms of R and δR as defined above, the inequality $E < E_r < E + \delta E$ becomes

$$R < \left(\sum_{i=1}^{3N} n_i{}^2\right)^{1/2} < R + \delta R.$$

Geometrically, this has the same meaning as before, except that now, our fictitious space is $3N$-dimensional. Nevertheless, it is clear that Ω must be equal to the
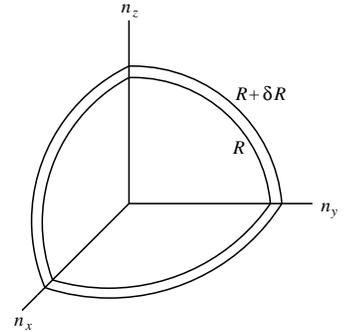
"volume" of a $3N$-dimensional shell of radius R and thickness δR, divided by a factor 2^{3N} to account for the fact that this space is semi-infinite.

We must therefore work out an expression for the "volume" of a $3N$-dimensional sphere. Surprisingly, this is not at all difficult, and it will be left as a homework problem. The desired expression is

$$V_{3N}(R) = \frac{\pi^{3N/2}}{\Gamma\left(\frac{3N}{2}+1\right)}\, R^{3N},$$

where $\Gamma(x)$ is the Gamma function, whose properties are explored in the same homework problem.

The volume of a $3N$-dimensional shell of radius R and thickness δR is then given by $\delta V_{3N} = V_{3N}(R + \delta R) - V_{3N}(R) = (dV_{3N}/dR)\delta R$, and we find $\Omega = 2^{-3N}\delta V_{3N}$. A few lines of algebra finally give

$$\boxed{\Omega(E, V, N; \delta E) = \frac{3N}{2}\, \frac{\pi^{3N/2}}{\Gamma\left(\frac{3N}{2}+1\right)}\, \frac{V^N}{(2\pi\hbar)^{3N}}\, (2mE)^{3N/2}\, \frac{\delta E}{E}\,.} \qquad (2.3)$$

Again, we find the same sort of expression as in Eq. (2.2).

Don't despair if you think that this calculation was just too much; this was probably the most challenging calculation of this entire course!

2.4 Statistical weight in classical mechanics

Our previous calculation of the statistical weight of an ideal gas was very much based upon quantum mechanics: the language of quantum mechanics was used to introduce the notion of a microstate, and the laws of quantum mechanics were used to count the number of microstates compatible with the given macroscopic description. However, the calculation also used the approximation of large quantum numbers, which allowed us to treat them as continuous variables, and to treat our fictitious $3N$-dimensional space as a continuum. But in this approximation, the energy eigenvalues are no longer quantized, but are continuous. Are we then not dealing with the *classical* world? According to Bohr's principle of correspondence, which states that a quantum system behaves classically when its quantum numbers are all very large, we are indeed. The question then arises as to whether we can derive Eq. (2.3) solely on the basis of classical mechanics.

The purpose of this section is to formulate the *classical* language of statistical mechanics, and to show how the classical rules provide another way of deriving Eq. (2.3).

2.4.1 Single particle in one dimension

We begin with the simplest example of a classical system, a single particle moving in one dimension. In classical mechanics, the *microstate* of the particle is specified by giving, at one particular time t, the value of both the particle's generalized coordinate q and its conjugate momentum p. The microstate at any other time can be obtained by integrating Hamilton's equations for the variables q and p. The q-p plane is called the *phase space* of the particle. A given microstate then corresponds to a particular point in phase space, and the motion of the particle traces a trajectory in phase space.

As a concrete example, consider a simple harmonic oscillator (SHO). Such a system is described by the Hamiltonian

$$H(p, q) = \frac{p^2}{2m} + \tfrac{1}{2}m\omega^2 q^2,$$

where m is the oscillator's mass, and ω its frequency. If the oscillator has energy E, then the trajectory it traces in phase space is the ellipse $H(p,q) = E$, or

$$1 = \frac{p^2}{2mE} + \frac{q^2}{2E/m\omega^2}.$$

This ellipse has semi-axes $\sqrt{2mE}$ and $\sqrt{2E/m\omega^2}$.

The *macrostate* of a classical system is specified in exactly the same way as for a quantum system, by giving the values of E, V, N, and δE. For our single one-dimensional particle, the macrostate is specified by giving the length of the "box" confining the particle, and by stating that the particle's Hamiltonian is restricted by

$$E < H(p,q) < E + \delta E.$$

This inequality defines a certain region of phase space; we call it the *allowed region*. For the SHO, this region is confined between the two ellipses at $H(p,q) = E$ and $H(p,q) = E + \delta E$.

In classical mechanics, the *statistical weight* is defined exactly as in quantum mechanics: It is the number of microstates compatible with the macroscopic description. Since microstates are here represented by points in phase space, what we must do is to count the number of phase-space points contained in the allowed region. In other words, we must calculate the *area* of the allowed region of phase space. This, however, cannot be our final result for Ω. This is because the area of phase space, unlike Ω, is not dimensionless: it carries the same dimension as $p \times q$, which has the dimension of an action, $\mathrm{kg\,m^2/s}$. To get a dimensionless number, we must divide the phase-space area by some constant h_0 which also carries that dimension. And if this theory is to have any kind of predictive power, this constant should be *universal*, that is, the same for all classical systems.

We therefore have the following prescription. In classical mechanics, the statistical weight is equal to the area of the allowed region of phase space, divided by the new universal constant h_0:

$$\Omega = (h_0)^{-1} \int_R dp\, dq.$$

Here, R designates the allowed region of phase space, which is defined by the condition $E < H(p,q) < E + \delta E$.

This prescription comes with a nice interpretation: We can imagine partitioning phase space into fundamental cells of area h_0; Ω is then just the total number of cells contained in the allowed region. The counting therefore introduces a fuzziness in phase space: two microstates are not distinguished if they are contained within the same cell. More precisely, for the purpose of computing the statistical weight, the microstates (p, q) and $(p + \Delta p, q + \Delta q)$ are considered to be different if and only if $\Delta p \Delta q > h_0$. This is highly reminiscent of the Heisenberg uncertainty principle, and indeed, we will see presently that h_0 has a lot to do with Planck's constant.

To elucidate the connection between h_0 and $\hbar$, we now compute Ω for a system consisting of a single SHO. We shall do this in two different ways: first classically, then quantum mechanically. In the classical calculation we must compute the area δA of the phase-space region between the two ellipses $H(p,q) = E$ and $H(p,q) = E + \delta E$. If we denote the area of the first ellipse by $A(E)$, and the area of the second ellipse by $A(E + \delta E)$, then $\delta A = A(E + \delta E) - A(E) = (dA/dE)\delta E$. Because the first ellipse has semi-axes $\sqrt{2mE}$ and $\sqrt{2E/m\omega^2}$, we have that $A(E) = \pi\sqrt{2mE}\sqrt{2E/m\omega^2} = 2\pi E/\omega$. It follows that $\delta A = 2\pi\delta E/\omega$. Finally, we obtain

$$\Omega = \frac{2\pi}{h_0}\frac{\delta E}{\omega}.$$
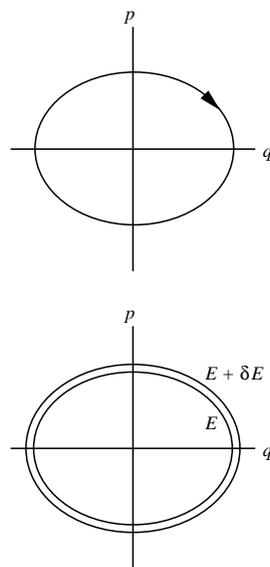
In the quantum calculation we must compute the number of eigenvalues $E_n = (n + \frac{1}{2})\hbar\omega \simeq n\hbar\omega$ contained in the interval between E and $E + \delta E$. (Recall that we work in the limit $n \gg 1$.) The condition for this is $E/\hbar\omega < n < E/\hbar\omega + \delta E/\hbar\omega$. Clearly, then,

$$\Omega = \frac{\delta E}{\hbar\omega}.$$

Both these expressions for Ω must be correct. It must therefore be that

$$h_0 \equiv 2\pi\hbar.$$

Because h_0 is a universal constant, this result is valid for *every* classical system.

2.4.2   General systems

It is easy to generalize this preceding prescription to systems possessing f degrees of freedom. Typically, $f = 3N \gg 1$. The general rules of classical statistical mechanics are as follows.

The microstate of an isolated, classical system is specified by giving, at some instant of time t, the value of the system's f generalized coordinates q_i, together with their conjugate momenta p_i. The $2f$-dimensional space of all q_i's and p_i's is called phase space.

The macrostate of the system is specified by giving the values of E, V, N, and δE. The condition

$$E < H(p_i, q_i) < E + \delta E$$

defines R, the allowed region of phase space. The statistical weight is then proportional to the "volume" of the allowed region:

$$\boxed{\Omega(E, V, N; \delta E) = (2\pi\hbar)^{-f} \int_R dp_1 \, dp_2 \, \cdots \, dp_f \, dq_1 \, dq_2 \, \cdots \, dq_f}. \qquad (2.4)$$

It is of course difficult to visualize the phase space when f is larger than 1. This, however, will not prevent us from doing calculations on the basis of Eq. (B.4).

2.4.3   Example: N particles in a box

We now show how to derive Eq. (2.3) using the classical definition of the statistical weight.

The Hamiltonian for a system of N noninteracting particles is

$$H(p_i, q_i) = \sum_{i=1}^{3N} \frac{p_i{}^2}{2m}.$$

Since H is actually independent of the coordinates, the condition $E < H(p_i, q_i) < E + \delta E$ constrains only the momenta. The coordinates, however, are constrained by the fact that the particles must all be confined to a box of volume V. This immediately gives us

$$\int dq_1 \, dq_2 \, \cdots \, dq_{3N} = V^N.$$

This follows because q_1, q_2, and q_3 can be chosen to be the x, y, and z coordinates of the first particle, and integrating over these clearly gives V.

The integral over the momenta is evaluated as follows. If we introduce new integration variables $\xi_i = p_i/\sqrt{2m}$, then

$$\int dp_1 \, dp_2 \, \cdots \, dp_{3N} = (2m)^{3N/2} \int d\xi_1 \, d\xi_2 \, \cdots \, d\xi_{3N}.$$

The domain of integration is defined by the inequality $E < H(p_i, q_i) < E + \delta E$, or

$$E < \sum_{i=1}^{3N} {\xi_i}^2 < E + \delta E.$$

Geometrically, this represents a $3N$-dimensional spherical shell of radius $\sqrt{E}$ and thickness $\sqrt{E + \delta E} - \sqrt{E} \simeq \delta E/2\sqrt{E}$. The integral over the ξ_i's is therefore equal to the "volume" of this spherical shell. (Since the ξ_i's can be negative as well as positive, there is no division by 2^{3N}.) But we know, from the work of Sec. 3, how to compute such a volume. After a bit of algebra, we obtain

$$\int dp_1 \, dp_2 \cdots dp_{3N} = \frac{3N}{2} \frac{\pi^{3N/2}}{\Gamma\left(\frac{3N}{2} + 1\right)} (2mE)^{3N/2} \frac{\delta E}{E}.$$

Finally, combining the results, we arrive at

$$\Omega(E, V, N; \delta E) = \frac{3N}{2} \frac{\pi^{3N/2}}{\Gamma\left(\frac{3N}{2} + 1\right)} \frac{V^N}{(2\pi\hbar)^{3N}} (2mE)^{3N/2} \frac{\delta E}{E},$$

which is exactly the same expression as in Eq. (2.3).

2.5   Fundamental postulates

Now that we have learned how to compute the statistical weight, both in quantum and classical mechanics, we may proceed with the formulation of the postulates of statistical mechanics for isolated systems.

An isolated system in a macrostate (E, V, N) can be in many different microstates, the number of which is given by the statistical weight Ω. What then can we say about the *probability* of finding the system in a *particular* quantum state ψ_r? Not much, in fact, and our first task in formulating the postulates of statistical mechanics is to admit our complete ignorance in this matter. There is just no way of knowing which set of quantum numbers will turn up when the microstate of the system is measured, and there would be no justification in assuming that some quantum states are preferred over others. This leads us to the *first postulate*:

> In *equilibrium*, an isolated system is *equally likely* to be found in *any* of its accessible microstates.

As a consequence of this postulate, we can immediately state that p_r, the probability of finding the system in the particular microstate ψ_r, must be given by

$$\boxed{p_r = \frac{1}{\Omega(E, V, N)}} . \tag{2.5}$$

Since the number of accessible microstates is Ω, the probabilities properly add up to one.

That the system must be in equilibrium for the first postulate to be true is extremely important: A system *out* of equilibrium is *not* equally likely to be found in any of its accessible microstates. To see this, consider the following situation. A gas is confined to one side (say the left side) of a box with the help of a removable partition. Before the partition is removed, the gas is in equilibrium, and all microstates are equally likely. We now imagine that the partition is suddenly removed. This clearly destroys the equilibrium, and the gas now starts filling the entire box, until a new equilibrium is established. But immediately after the partition is removed, the gas is *still* found on the left side, even though the entire box is now accessible
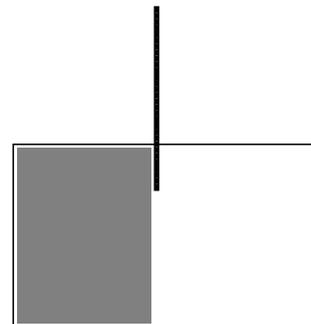
to it. This is just the statement that some microstates are more likely than others: Microstates corresponding to finding the gas on the left side are much more likely than those corresponding to finding the gas on the right side. This is sufficient to conclude that equilibrium is required for the microstates to be equally likely. Out of equilibrium, some microstates are more likely than others.

We may take these ideas a step further. It should be clear that when it is out of equilibrium, the system is *always* found in a *subset* of all its accessible microstates. (Immediately after the partition is removed, the gas *will* be found on the left side, a subset of the entire box.) It should also be clear that the number of microstates in this subset is *smaller* than the total number of microstates accessible to the system. Now, as the system proceeds toward equilibrium, the subset becomes larger and larger, until it eventually becomes the *entire set* of accessible microstates; at this point the system has found a new equilibrium. It is therefore at equilibrium that the system can be found in the *largest* number of microstates. Just like the entropy, this number never decreases, and is maximum at equilibrium. This strongly suggests that S must be closely related to Ω.

To determine what this relation is, we consider a composite system C consisting of two parts, A and B. We assume that this system is isolated and in equilibrium. We want to express Ω_C, the number of microstates accessible to C, in terms of Ω_A and Ω_B. We denote by $\psi_r(A)$ the quantum states accessible to A, and by $\psi_s(B)$ those accessible to B. The states accessible to C can then be written as the product $\psi_r(A)\psi_s(B)$. How many of these states are there? Well, since r takes a number Ω_A of values, while s takes a number Ω_B of values, and since any value of r can be combined with any value of s to form one of C's microstates, the answer is clearly

$$\Omega_C = \Omega_A \Omega_B.$$

The statistical weights *multiply*. The entropies, on the other hand, *add*:

$$S_C = S_A + S_B.$$

So if S is a function of Ω, this function must be such that when two statistical weights are multiplied, the resulting entropies add. There is only one such function: the logarithm. So it must be that $S \propto \ln \Omega$, and the constant of proportionality is *defined* to be Boltzmann's constant k.

We have obtained our *second postulate*:

> The entropy of an isolated system in equilibrium is related to its statistical weight by the relation
>
> $$\boxed{S(E, V, N) = k \ln \Omega(E, V, N)} \,, \qquad (2.6)$$
>
> where k is Boltzmann's constant.

This definition of the entropy automatically provides a *proof* of the third law of thermodynamics, which states that a system at zero temperature must have zero entropy. Indeed, a system at zero temperature must necessarily be in its ground state, because thermal excitations to other states are impossible. So the number of accessible states is exactly one, and Eq. (2.6) implies that the entropy vanishes.

The usefulness of the second postulate resides mostly in the fact that it permits the computation of *all* of the system's thermodynamic quantities. To see this, we only need write the first law of thermodynamics in the form

$$dS = \frac{1}{T}\, dE + \frac{P}{T}\, dV - \frac{\mu}{T}\, dN,$$

and recall that this equation implies the following definitions for T, P, and μ:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N}, \qquad \frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_{E,N}, \qquad \frac{\mu}{T} = -\left(\frac{\partial S}{\partial N}\right)_{E,V}. \tag{2.7}$$

When we first encountered such relations, in part A of this course, they were just formal relations: The equation for dS implied the existence of a function $S(E, V, N)$, but it gave us no way of actually computing this function; and although T, P, and μ could be calculated from this function, this could not be done in practice because the function was not known.

Things are different now, because the second postulate gives us a way of actually computing the function $S(E, V, N)$. All we have to do is first calculate $\Omega(E, V, N)$, and then take a logarithm. And once the function $S(E, V, N)$ is known, the system's thermodynamics can be determined from the relations (2.7).

2.6   Application: ideal gas

2.6.1   Entropy

In Sec. 3 (and also in Sec. 4), we worked out an expression for the statistical weight of a system of N noninteracting particles, which is precisely what we mean by an "ideal gas". The result is Eq. (2.3), above. To compute the entropy we must take the natural logarithm of Ω, which is

$$\ln \Omega = \ln\left(\frac{3N}{2}\right) - \ln\left[\Gamma\left(\frac{3N}{2} + 1\right)\right] + N \ln\left[\frac{V}{(2\pi\hbar)^3}\,(2\pi m E)^{3/2}\right] + \ln\left(\frac{\delta E}{E}\right).$$

To simplify the second term on the right-hand side, we make use of the Stirling approximation, which states that for large numbers n,

$$\ln\left[\Gamma(n+1)\right] = \ln(n!) \simeq n \ln n - n.$$

The number $3N/2$ is certainly very large, and our expression for $\ln \Omega$ becomes

$$\ln \Omega = N \ln\left[\frac{V}{(2\pi\hbar)^3}\left(\frac{4\pi m E}{3N}\right)^{3/2}\right] + \frac{3N}{2} + \ln\left(\frac{3N}{2}\right) + \ln\left(\frac{\delta E}{E}\right).$$

This can be simplified further, because $\ln(3N/2)$ can be neglected in front of $3N/2$; and since $\ln(\delta E/E)$ is now the only term that is not proportional to N, it too can be neglected.

We finally arrive at the expression

$$S = Nk \ln\left[\frac{V}{(2\pi\hbar)^3}\left(\frac{4\pi m E}{3N}\right)^{3/2}\right] + \frac{3Nk}{2}$$

for the entropy of an ideal gas. This, however, is obviously a wrong result!

2.6.2   Gibbs' paradox

Why is this obviously wrong? The answer is that S should be an extensive variable. If $N \to \lambda N$, $V \to \lambda V$, and $E \to \lambda E$, then it should also be that $S \to \lambda S$. But this is not what we find here. Instead, we find $S \to \lambda S + \lambda Nk \ln \lambda$, which is *not* the correct scaling behaviour. This puzzling result is known as *Gibbs' paradox*.

We have made a mistake in our calculation. What went wrong? The problem is that when computing Ω in Sec. 3, we have actually *over-counted* the number

of microstates by implicitly assuming that the particles were distinguishable. This was wrong: in quantum mechanics, identical particles are *indistinguishable*; we must account for this in our calculation.

When did we wrongly assume that the particles were distinguishable? This we did in the very way we set up the problem. We imagined that each of the N particles carried three quantum numbers (n_x, n_y, and n_z) that it could call its own. If we group these numbers into a vector $\boldsymbol{n} = (n_x, n_y, n_z)$, then we can say that particle a has quantum numbers $\boldsymbol{n}_a$, while particle b has quantum numbers $\boldsymbol{n}_b$. However, we must ask: Is the state obtained by interchanging the particles a and b actually *distinct* from the original state? [For concreteness, suppose that $\boldsymbol{n}_a = (2, 9, 4)$ and $\boldsymbol{n}_b = (5, 2, 7)$; is this state distinct from the state $\boldsymbol{n}_a = (5, 2, 7)$ and $\boldsymbol{n}_b = (2, 9, 4)$?] We have treated them as distinct in our calculation. But according to the postulates of quantum mechanics, these states are *not* distinct; they are one and the same, as would also be the states obtained by the interchange of *any* number of particles. Identical particles are *not* distinguishable in quantum mechanics, and it is just wrong to think of them as carrying their own individual quantum numbers.

Fortunately, it is not too difficult to take care of the over-counting, and thus correct our previous expression for Ω. Clearly, what we must do is divide Ω by the total number of possible interchanges among the N particles. In other words, we must divide Ω by the total number of *permutations* of the N particles among themselves. This number is equal to $N!$, and

$$\boxed{\Omega_{\text{correct}} = \frac{\Omega_{\text{wrong}}}{N!}} . \tag{2.8}$$

We can immediately use this to correct our previous result for the entropy. We have $S_{\text{correct}} = S_{\text{wrong}} - k \ln(N!) = S_{\text{wrong}} - kN \ln N + kN$, after using the Stirling approximation. Our final result is therefore

$$\boxed{S = Nk \ln \left[\frac{V}{(2\pi\hbar)^3 N} \left(\frac{4\pi mE}{3N} \right)^{3/2} \right] + \frac{5Nk}{2}} . \tag{2.9}$$

The entropy is now properly extensive.

2.6.3   Thermodynamic quantities

Now that we have the correct expression for the entropy of an ideal gas, it is a simple matter to use Eqs. (2.7) to derive expressions for the other thermodynamic variables.

The temperature is defined by $1/T = \partial S/\partial E$. Since $S = \frac{3}{2} Nk \ln E$ plus terms that don't involve E, we obtain $1/T = \frac{3}{2} Nk/E$, or

$$\boxed{E = \frac{3}{2} NkT} . \tag{2.10}$$

This equation tells us that the mean energy per particle is $\frac{3}{2}kT$, or $\frac{1}{2}kT$ for each degrees of freedom.

On the other hand, the pressure is defined by $P/T = \partial S/\partial V$. Since $S = Nk \ln V$ plus terms that don't involve V, we obtain $P/T = Nk/V$, or

$$\boxed{PV = NkT} . \tag{2.11}$$

This is the familiar equation of state for an ideal gas, which was used many times in part A of this course. The distinction between the two contexts should be clear: In

thermodynamics, the equation of state must be provided as input, and cannot be derived; in statistical mechanics, however, *the equation of state can be derived from first principles.* This derivation involves first the formulation of a physical model for the system (an ideal gas consists of noninteracting particles), and then the calculation of the statistical weight. Everything else follows with simple manipulations (taking a logarithm, and evaluating a few partial derivatives).

We shall leave the evaluation of μ, the chemical potential, as a homework problem.

2.7 Problems

1. (Reif 1.24)

   a) A particle is equally likely to lie anywhere on the circumference of a circle. Consider as the z axis any straight line in the plane of the circle and passing through its centre. Denote by θ the angle between this z axis and the straight line connecting the centre to the particle. What is the probability that this angle lies between θ and $\theta + d\theta$?

   b) A particle is equally likely to lie anywhere on the surface of a sphere. Consider any line through the centre of this sphere as the z axis. Denote by θ the angle between this z axis and the straight line connecting the centre of the sphere to the particle. What is the probability that this angle lies between θ and $\theta + d\theta$?

2. The Gaussian probability distribution function is given by

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right].$$

   a) Show that $p(x)$ is correctly normalized.

   b) Show that

$$\overline{(x-\mu)^n} = \frac{2^{\frac{n}{2}}}{\sqrt{\pi}}\Gamma\left(\frac{n+1}{2}\right)\sigma^n$$

   if n is even, while $\overline{(x-\mu)^n} = 0$ if n is odd. Then use the properties of the Γ function to show that for n even, the previous equation reduces to

$$\overline{(x-\mu)^n} = (n-1)!!\,\sigma^n,$$

   where $(n-1)!! \equiv (n-1)\cdot(n-3)\cdot(n-5)\cdots 3\cdot 1$. What are $\bar{x}$ and $\overline{(x-\bar{x})^2}$?

3. This problem is an exercise in pure mathematics. We intend to show that the volume of an n-dimensional "sphere" of radius R is given by

$$V_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}R^n, \tag{2.12}$$

   where $\Gamma(x)$ denotes the Gamma function. This result is used in text. To derive this we will need the integral

$$I_n \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left[-\sum_{i=1}^{n}x_i^2\right]dx_1 dx_2\cdots dx_n. \tag{2.13}$$

a) The first step is to get familiar with the Gamma function, if you are not already. For any real number $x > 0$, the Gamma function is defined by

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} \, dt.$$

By integrating by parts, show that $\Gamma(x + 1) = x\Gamma(x)$. Show also that $\Gamma(1) = 1$. Use these results to show that $\Gamma(m+1) = m!$ if m is an integer.

b) Argue that V_n must take the form $V_n = C_n R^n$, where C_n is a constant to be determined, and $R^2 = \sum_{i=1}^n x_i^2$ is the n-dimensional squared radius. Then argue that I_n can be written as $I_n = \int e^{-R^2} dV_n$, which is a simple integration over the single variable R. (This step is trickier than it seems. Don't be fooled by notation: dV_n is the difference in the volumes of two concentric spheres, one of radius R, the other of radius $R+dR$; dV_n is *not equal* to $dx_1 \cdots dx_n$.) Combining all this, show that $I_n = C_n \Gamma(\frac{n}{2} + 1)$.

c) Now go back to Eq. (2.13). Show that $I_n = (I_1)^n$. Of all these integrals, $I_2 = \int_{-\infty}^\infty e^{-(x^2+y^2)} \, dxdy$ is easiest to evaluate. By going to polar coordinates, show that $I_2 = \pi$. So $I_n = \pi^{\frac{n}{2}}$.

d) Combining the results of parts b) and c), show that $C_n = \pi^{\frac{n}{2}}/\Gamma(\frac{n}{2} + 1)$. This then proves Eq. (2.12).

e) Show that $\Gamma(\frac{1}{2}) = I_1 = \sqrt{\pi}$. Use this and Eq. (2.12) to show that $V_3 = 4\pi R^3/3$, a familiar result.

4. In this problem we study a system consisting of N elements (let's call them spins), each capable of having the values $\pm\varepsilon$ for its energy. Each spin is imagined to be located on a particular site on a lattice. Denoting by n the number of spins with energy $+\varepsilon$, we have that the total energy E is given by $E = n\varepsilon + (N - n)(-\varepsilon)$, or

$$E = (2n - N)\varepsilon.$$

a) Calculate $\Omega(E)$, the number of ways the system can be prepared so that its energy will have the value E. (Hint: to have energy E the system must have n spins with energy ε and $N - n$ spins with energy $-\varepsilon$. These spins can be put on the lattice in any order order whatsoever.) Express your result in terms of E, N, and ε.

b) Calculate the entropy of the system. To simplify the expressions, assume that N, $N + E/\varepsilon$, and $N - E/\varepsilon$ are all very large, and use the Stirling approximation $\ln(M!) = M \ln M - M$ for large M.

c) Calculate the temperature. When can the temperature become negative? (The physical meaning of negative temperatures will be discussed later on in the course.)

d) Show that the energy can be expressed as

$$E = -N\varepsilon \tanh(\varepsilon/kT).$$

How does the energy behave at low (positive) temperatures? Give a physical explanation for the result. How does the energy behave at high (positive) temperatures? Again give a physical explanation for the result. Can the energy ever be positive?

5. (Reif 3.5) A system consists of N_1 molecules of type 1 and N_2 molecules of type 2 confined within a box of volume V. The molecules are supposed to interact very weakly so that they constitute an ideal gas mixture.

a) How does the total number $\Omega(E, V, N_1, N_2)$ of microstates in the range between E and $E + \delta E$ depend on V? You may treat the problem classically.

b) Use this result to find the equation of state of this system, i.e., to find its pressure P as a function of V, T, N_1, and N_2.

6. We have seen that the statistical weight $\Omega(E, V, N)$, as calculated using classical mechanics, is proportional to the "volume" of the region of phase space accessible to the system:

$$\Omega(E, V, N) \propto \int dp_1 \cdots dp_f \, dq_1 \cdots dq_f.$$

For this to make sense, it must be true that $\Omega(E, V, N)$ is the same no matter what coordinate system is used to describe the system; we must be able to use any set of generalized coordinates q_i, together with the corresponding momenta p_i. It must then be true that the "volume element" $dp_1 \cdots dp_f \, dq_1 \cdots dq_f$ is invariant under a coordinate transformation $q_i \to q'_i$, together with the corresponding transformation on the momenta, $p_i \to p'_i$. That this indeed the case is known as Liouville's theorem. (This theorem also holds for the wider class of transformations known as *canonical transformations*.)

Let us consider an example in order to convince ourselves that Liouville's theorem really holds. Consider a particle moving in the xy plane. Using Cartesian coordinates, the Lagrangian is

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, y),$$

so that the momenta are $p_x = m\dot{x}$ and $p_y = m\dot{y}$.

a) The motion of the particle can equally well be described using polar coordinates. Show that in this case the Lagrangian is

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r, \theta),$$

and that the momenta are $p_r = m\dot{r}$ and $p_\theta = mr^2\dot{\theta}$.

b) Write down the transformation of the phase space coordinates when one goes from Cartesian to polar coordinates. In other words, express p_x, p_y, x, and y in terms of p_r, p_θ, r, and θ.

c) Now show that the volume elements $dp_x dp_y dx dy$ and $dp_r dp_\theta dr d\theta$ are equal. (Hint: What is the Jacobian associated with the above transformation of phase-space coordinates?)

d) What would happen if the state of a system were specified by the coordinates q_i and the velocities $\dot{q}_i$, instead of the coordinates and the momenta? (Hint: compare the volume elements $d\dot{x}d\dot{y}dxdy$ and $d\dot{r}d\dot{\theta}drd\theta$.)

7. We have seen that the entropy of an isolated ideal gas is given by

$$S = Nk \, \ln\left[\frac{V}{(2\pi\hbar)^3 N}\left(\frac{4\pi mE}{3N}\right)^{3/2}\right] + \frac{5Nk}{2}.$$

a) Calculate the chemical potential μ. Express your result in terms of T, V, and N.

b) Give the numerical value of $\alpha \equiv -\mu/kT$ for a Helium gas at room temperature and at atmospheric pressure.

c) In the light of the definition

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{S,V},$$

explain how μ can be negative.

8. Calculate the statistical weight of a *classical* system consisting of N one-dimensional simple harmonic oscillators. The system is assumed to be in isolation, with an energy in the interval between E and $E+\delta E$, where $\delta E \ll E$. The system's Hamiltonian is

$$H(p_i, q_i) = \sum_{i=1}^{N} \left(\frac{p_i{}^2}{2m} + \frac{1}{2}m\omega^2\, q_i{}^2\right).$$

You may assume that the oscillators are distinguishable, so that the statistical weight is given by

$$\Omega(E, N; \delta E) = (2\pi\hbar)^{-N} \int_R dp_1 \cdots dp_N\, dq_1 \cdots dq_N,$$

where R is the allowed region of phase space.

CHAPTER 3
STATISTICAL MECHANICS OF INTERACTING SYSTEMS

3.1 System in contact with a reservoir

In part B of this course, the rules of statistical mechanics were formulated for the restricted case of a system in complete isolation. Here we generalize the formulation to the more common situation in which the system is interacting with a *reservoir*. During this interaction, the system's energy is allowed to change, and so is its number of particles; its volume, however, will be assumed to be fixed at all times.

3.1.1 Probability distributions

The first thing we would like to calculate is p_r, the probability of finding the interacting system in the particular microstate ψ_r. For isolated systems we found that the probabilities were all equal to $1/\Omega$, the inverse of the statistical weight; this is not true for interacting systems.

To find p_r, we consider a system A in contact with a reservoir R at temperature T, pressure P, and chemical potential μ. (Actually, because we keep the volumes of A and R fixed, P will not be a relevant quantity in this analysis.) The combined system $C = A + R$ is assumed to be isolated and in equilibrium; the previous postulates of statistical mechanics therefore apply to it. In a general interaction between A and R, both heat and particles are allowed to flow back and forth, but as was already stated, the respective volumes are kept fixed.

At any moment of time the system A has energy E_A, number of particles N_A, statistical weight Ω_A, and entropy $S_A = k \ln \Omega_A$. Similarly, at any moment of time the reservoir R has energy E_R, number of particles N_R, statistical weight Ω_R, and entropy $S_R = k \ln \Omega_R$. These quantities are all susceptible to change during the interaction. However, since the combined system is isolated and in equilibrium, the quantities $E_C = E_A + E_B$, $N_C = N_A + N_B$, $\Omega_C = \Omega_A \Omega_B$, and $S_C = S_A + S_B$ are all constant. We assume that the reservoir R is much larger than the system A, so that

$$E_A \ll E_C, \qquad N_A \ll N_C.$$

At any moment of time the system A can be in any one of its accessible microstates $\psi_r(A)$, but these states are not all equally likely. Similarly, at any moment of time the reservoir R can be in any one of its accessible microstates $\psi_s(R)$, but these states are also not all equally likely. The states of the combined system C will be denoted $\psi(C) = \psi_r(A)\psi_s(R)$; because C is isolated, these states *are* all equally likely.
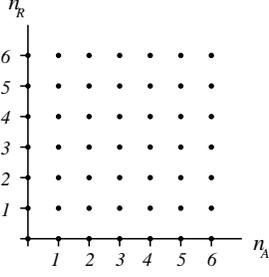
To make these ideas more concrete, we consider the following example: We suppose that both A and R are a single simple harmonic oscillator. The state of A is specified by giving the value of n_A, its quantum number, and similarly, the state of R is specified by giving the value of n_R. The state of the combined system C is then specified by giving the values of both n_A and n_R. The states accessible to C can therefore be represented by points in the n_A-n_R plane.

We want to find p_{r^*}, the probability that A will be found in the particular microstate $\psi_{r^*}(A)$. [Here, r^* denotes a *particular value* of the index r, which labels all the microstates accessible to A. In other words, the microstate $\psi_{r^*}(A)$ is just *one* of all the microstates accessible to A.] This calculation requires several steps.

First, we suppose that A *actually is* in the particular state ψ_{r^*}. This implies that $E_A = E_{r^*}$, where E_{r^*} is the energy eigenvalue of this state, and $N_A = N_{r^*}$, where N_{r^*} is the number of particles contained in the state ψ_{r^*}. It follows that for the reservoir R, we must have $E_R = E_C - E_{r^*}$ and $N_R = N_C - N_{r^*}$.

Second, we notice that choosing the particular state ψ_{r^*} for A automatically divides the states accessible to C into two separate groups. The first group consists of all states $\psi(C)$ such that A is in the state ψ_{r^*}; these states can be denoted $\psi_{r=r^*}(C)$, and are of the form $\psi_{r^*}(A)\psi_s(B)$. The second group consists of all states $\psi(C)$ such that A is *not* in the state ψ_{r^*}; these can be denoted $\psi_{r\neq r^*}(C)$, and are of the form $\psi_r(A)\psi_s(B)$, where r is any number other than r^*. In our example above, the first group consists of those states for which n_A has a specific value, say $n_A = 4$; this group can be represented by a vertical line in the n_A-n_R plane. The second group consists of all other states.

Third, we calculate $\Omega_C^{r=r^*}$, the number of states contained in the first group. To do this, we use the identity $\Omega_C = \Omega_A\Omega_R$. Since the number of states accessible to A is equal to one when $\psi(C)$ belongs to the first group, we have that $\Omega_C = \Omega_R$, or

$$\Omega_C^{r=r^*} = \Omega_R\big(E_C - E_{r^*}, N_C - N_{r^*}\big),$$

where we have indicated the dependence of Ω_R on E_R and N_R. The number of states contained in the second group is then obviously $\Omega_C^{r\neq r^*} = \Omega_C - \Omega_C^{r=r^*}$.

Fourth, we calculate the probability that the combined system C will be found in a state $\psi(C)$ belonging to the first group. This step is easy. Since the states accessible to C are all equally likely, this probability must be equal to the ratio $\Omega_C^{r=r^*}/\Omega_C$. (This is a consequence of the "or" rule.) Similarly, the probability that C will be found in a state belonging to the second group is $\Omega_C^{r\neq r^*}/\Omega_C$, and the probabilities correctly add up to one. We therefore have

$$P\big[\psi_{r=r^*}(C)\big] = \frac{\Omega_R(E_C - E_{r^*}, N_C - N_{r^*})}{\Omega_C}.$$

Fifth, and finally, we relate this to p_{r^*}, the probability of finding A in the state ψ_{r^*}. This step is also easy. Of all the states accessible to C, the probability of selecting a member of the first group is $P[\psi_{r=r^*}(C)]$, which we have just calculated. But once we have selected such a state, the probability of finding A in the state ψ_{r^*} is unity, because *all* states in the first group are of the form $\psi_{r^*}(A)\psi_s(B)$. The overall probability of finding A in the state $r = r^*$ is therefore *precisely equal* to $P[\psi_{r=r^*}(C)]$. The desired result is therefore

$$p_{r^*} = \frac{\Omega_R(E_C - E_{r^*}, N_C - N_{r^*})}{\Omega_C}.$$

This result holds irrespective of what the state ψ_{r^*} is chosen to be, since this was left arbitrary from the beginning. The result must therefore hold for *any* of the states accessible to A. And since the statistical weights can be related to the entropies,
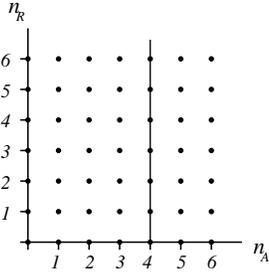
we shall express our probabilities as

$$p_r = \exp\left(\frac{S_R(E_C - E_r, N_C - N_r) - S_C}{k}\right).\tag{3.1}$$

Thus far this discussion has been quite general, and we have not yet used the assumption that the reservoir R is much larger than the system A. But now we use the relations $E_r \ll E_C$ and $N_r \ll N_C$ in order to simplify our expression (3.1) for the probabilities. To do so, we simply expand $S_R(E_R, N_R)$ in a Taylor series about (E_C, N_C), which yields

$$S_R(E_C - E_r, N_C - N_r) \simeq S_R(E_C, N_C) + \frac{\partial S_R}{\partial E_R}(-E_r) + \frac{\partial S_R}{\partial N_R}(-N_r).$$

Using the definitions $1/T = \partial S_R/\partial E_R$ and $\mu/T = -\partial S_R/\partial N_R$, and the fact that E_C, N_C, and S_C are just constants, Eq. (3.1) becomes

$$p_r \propto \exp\left[-(E_r - \mu N_r)/kT\right].$$

The constant of proportionality is simply determined by the requirement that the probabilities must all add up to one: $\sum_r p_r = 1$. This gives our final result,

$$p_r = \frac{1}{\mathcal{Z}}\, e^{-(E_r - \mu N_r)/kT}, \qquad \mathcal{Z} = \sum_r e^{-(E_r - \mu N_r)/kT}.\tag{3.2}$$

This distribution of probabilities is known as the *Gibbs distribution*, and the quantity $\mathcal{Z}$ is called the *grand partition function*; it is obtained by summing the Gibbs factor $\exp[-(E_r - \mu N_r)/kT]$ over *all* states accessible to the system A. Because of the interaction between A and the reservoir, these states are not restricted to all have the same value for the energy and the number of particles, as is the case for isolated systems.

For the special case where the interaction does not allow an exchange of particles, Eq. (3.2) simplifies. Since the number of particles is now the same for all states, $N_r = N =$ constant, and we have

$$p_r = \frac{e^{\mu N/kT}}{\mathcal{Z}}\, e^{-E_r/kT}.$$

In other words,

$$p_r = \frac{1}{Z}\, e^{-E_r/kT}, \qquad Z = \sum_r e^{-E_r/kT}.\tag{3.3}$$

This is the *Boltzmann distribution*, and Z, the *partition function*, is obtained by summing the Boltzmann factor $\exp(-E_r/kT)$ over *all* states accessible to the system.

3.1.2   Entropy

Now that we have the probability distributions, what can we say about the system's entropy? We recall from Eq. (3.1) that

$$-k \ln p_r = S_C - S_R(E_C - E_r, N_C - N_r) = S_A(E_r, N_r),$$

and it seems that we have our answer: $-k \ln p_r$ is the system's entropy *if* it happens to be in the microstate ψ_r. But if A is *known* to be in this state, then $p_r = 1$, and the entropy is zero. So this equation is not quite what we want!

Of course, the whole purpose of statistical mechanics is to provide a way of describing the macroscopic properties of a system whose microstate is not only unknown, but also susceptible to change because of the interaction with the reservoir. Since the microstate is unknown, and because we want the entropy to be a *macroscopic* quantity, it makes sense to *average* the microscopic entropy $-k \ln p_r$ over all the microstates that are accessible to the system. We therefore *define* the macroscopic entropy to be

$$S = \overline{S(E_r, N_r)} = -k \, \overline{\ln p_r}.$$

The entropy of an interacting system is therefore given by

$$\boxed{S = -k \sum_r p_r \ln p_r} . \tag{3.4}$$

This expression holds both for the Gibbs and the Boltzmann distributions.

In fact, since an isolated system is just a special case of an interacting one, Eq. (3.4) should apply for this case as well. It is easy to see that this is indeed true: For isolated systems, $p_r = 1/\Omega$, and $S = -k \sum_r (1/\Omega) \ln(1/\Omega) = k \ln \Omega \sum_r (1/\Omega)$. Since the sum contains a number Ω of equal terms, we arrive at

$$S = k \ln \Omega,$$

which is the correct answer. We therefore conclude that Eq. (3.4) is universally valid.

3.2   Boltzmann distribution

In this section we consider a system in thermal contact with a heat reservoir at temperature T, whose statistical mechanics is governed by the Boltzmann distribution (3.3). It is convenient to introduce the inverse-temperature parameter

$$\boxed{\beta = \frac{1}{kT}} . \tag{3.5}$$

In terms of this, the Boltzmann distribution reads

$$\boxed{p_r = \frac{1}{Z} e^{-\beta E_r}, \qquad Z = \sum_r e^{-\beta E_r}} . \tag{3.6}$$

We ask: how do we calculate the system's various thermodynamic quantities, such as internal energy, entropy, pressure, and chemical potential?

3.2.1   Thermodynamic quantities

Since the system is not in a state of definite energy, because of its interaction with the reservoir, the macroscopic quantity E must represent the system's *mean internal energy*, just as the macroscopic quantity S was earlier defined to be the system's mean entropy. So we define

$$E \equiv \overline{E_r} = \sum_r p_r E_r = \frac{1}{Z} \sum_r E_r e^{-\beta E_r}.$$

However,

$$-\frac{\partial Z}{\partial \beta} = -\sum_r (-E_r) e^{-\beta E_r} = ZE,$$

and we obtain the useful result

$$E = -\frac{\partial \ln Z}{\partial \beta}.$$ (3.7)

So, if the partition function can be evaluated as a function of β, then the system's mean energy can be obtained immediately by partial differentiation. This will be a recurring theme in what follows.

The system's macroscopic entropy is defined by Eq. (3.4). To compute it efficiently, we express $\ln p_r$ as $-(\ln Z + \beta E_r)$, so that

$$S = k \sum_r p_r (\ln Z + \beta E_r).$$

The first term on the right-hand side is equal to $k \ln Z \sum_r p_r = k \ln Z$, since the probabilities are normalized. The second term, on the other hand, is equal to $k\beta E$. This gives

$$S = k(\ln Z + \beta E)$$ (3.8)

for the entropy of a system interacting with a heat reservoir.

Equation (3.8) is extremely important, for it implies that the Helmholtz free energy $F = E - TS$ is directly related to the partition function. Indeed, a few steps of algebra yield

$$F(T, V, N) = -kT \ln Z(T, V, N).$$ (3.9)

That F, and therefore also Z, depend on the quantities T, V, and N follows from the relation

$$dF = -S\, dT - P\, dV + \mu\, dN,$$

derived in part A of this course (Sec. A7). This relation also gives definitions for S, P, and μ in terms of partial derivatives of F:

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \qquad P = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \qquad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V}.$$ (3.10)

Notice that the equation $S = -\partial F / \partial T$ is exactly the same equation as Eq. (3.8).

So, if the partition function is known as a function of T, V, and N, then the Helmholtz free energy is known, and *all other* thermodynamic quantities can be obtained by partial differentiation. In other words:

> For a system in thermal contact with a heat reservoir at a temperature T, every thermodynamic quantity can be calculated directly from the partition function $Z(T, V, N)$.

This is highly analogous to the situation holding for isolated systems, where every thermodynamic quantity can be calculated from the statistical weight $\Omega(E, V, N)$.

The key to the thermodynamic quantities is therefore the evaluation of the partition function. That this quantity must depend on T is obvious, because $\beta = 1/kT$ appears explicitly in the Boltzmann factor. The origin of the V and N dependence is more subtle. The partition function depends on V because E_r, the energy eigenvalues, usually depend on V. (As an example, the energy eigenvalues of a system of noninteracting particles are proportional to $V^{-2/3}$.) The partition function also depends on N because $\sum_r$, the sum over states, depends on N: the larger the number of particles, the larger the number of states accessible to the system.

3.2.2 Energy distribution

We have associated E, the thermodynamic energy, with the *averaged* value of the system's energy. However, when the energy is actually *measured*, the laws of quantum mechanics guarantee that one of the energy eigenvalues will be returned. (Which eigenvalue it will be cannot be predicted with certainty, because the system's quantum state keeps on changing as a result of the interaction with the reservoir.) In general, E is *not equal* to any one of the energy eigenvalues. The question therefore arises as to how strongly the mean value E candeviate, at any given time, from the actual energy of the system. The answer, as we shall see, is not very much. For all practical purposes, at any given time the mean energy is identical to the system's actual energy.

To analyze this question, we consider the quantity $p(\mathcal{E})\,d\mathcal{E}$, the probability that the system's energy will be measured to be in the interval between $\mathcal{E}$ and $\mathcal{E} + d\mathcal{E}$. This quantity is computed as follows. By using the "or" rule, we can immediately write that $p(\mathcal{E})\,d\mathcal{E} = \sum_r' p_r$, where the sum extends to all microstates whose energy eigenvalues are in the specified range. Since these states have (more or less) the same energy, this gives $p(\mathcal{E})\,d\mathcal{E} = Z^{-1}\exp(-\beta\mathcal{E})(\sum_r' 1)$. Now, the number of terms in the sum is equal to the number of states whose energy eigenvalues are in the specified range. But this is just the definition of the statistical weight, $\Omega(\mathcal{E}; d\mathcal{E}) = \omega(\mathcal{E})d\mathcal{E}$. We therefore obtain

$$\boxed{p(\mathcal{E})\,d\mathcal{E} = \frac{1}{Z}\,\omega(\mathcal{E})e^{-\beta\mathcal{E}}\,d\mathcal{E}, \qquad Z = \int \omega(\mathcal{E})e^{-\beta\mathcal{E}}\,d\mathcal{E}}, \qquad (3.11)$$

where $\omega(\mathcal{E})$ is the density of states. Once again the partition function has been defined so as to make the probabilities add up to one.

The probability distribution function for the energy can now be used to answer our question. This function is the product of two functions of $\mathcal{E}$, the density of states $\omega(\mathcal{E})$ and the Boltzmann factor $\exp(-\beta\mathcal{E})$. Typically, the density of states increases very rapidly with the energy. (As an example, we may recall our result for the ideal gas, $\omega \propto \mathcal{E}^{3N/2}$. For $N \sim 10^{23}$, this really does increase rapidly!) The Boltzmann factor, on the other hand, decreases very rapidly (exponentially!). It follows that the *product* of these two functions will be almost zero almost everywhere, *except* for a very narrow region near $\mathcal{E} = E$, the mean value. In other words, $p(\mathcal{E})$ is very sharply peaked about E, and has a very small standard deviation.

We now put this on a quantitative basis by computing ΔE, the energy's root-mean-square fluctuation. This is defined by

$$\boxed{(\Delta E)^2 = \overline{(\mathcal{E} - E)^2}}, \qquad (3.12)$$

where $E \equiv \overline{\mathcal{E}}$. To calculate ΔE, we first notice that Eq. (3.12) implies

$$(\Delta E)^2 = \overline{\mathcal{E}^2} - \overline{\mathcal{E}}^2.$$

Next, we use the following expression for the mean energy,

$$E = \int \mathcal{E}p(\mathcal{E})\,d\mathcal{E} = \frac{\int \mathcal{E}\omega(\mathcal{E})e^{-\beta\mathcal{E}}\,d\mathcal{E}}{\int \omega(\mathcal{E})e^{-\beta\mathcal{E}}\,d\mathcal{E}},$$

and compute

$$-\frac{\partial E}{\partial \beta} = \frac{\int \mathcal{E}^2\omega(\mathcal{E})e^{-\beta\mathcal{E}}\,d\mathcal{E}}{\int \omega(\mathcal{E})e^{-\beta\mathcal{E}}\,d\mathcal{E}} - \left(\frac{\int \mathcal{E}\omega(\mathcal{E})e^{-\beta\mathcal{E}}\,d\mathcal{E}}{\int \omega(\mathcal{E})e^{-\beta\mathcal{E}}\,d\mathcal{E}}\right)^2.$$

Finally, we recognize in the first term of the right-hand side the definition for $\overline{\mathcal{E}^2}$, while the second term gives $\overline{\mathcal{E}}^2$. We therefore have obtained

$$\boxed{(\Delta E)^2 = -\left(\frac{\partial E}{\partial \beta}\right)_{V,N}} \,, \tag{3.13}$$

which tells us that the energy fluctuation can be computed directly from the energy itself, by taking a partial derivative.

As a concrete example, let us consider the ideal gas, for which $E = \frac{3}{2}NkT = \frac{3}{2}N/\beta$. We immediately obtain $(\Delta E)^2 = \frac{3}{2}N/\beta^2$. To express this in a meaningful way, we should compare ΔE to E itself. A few steps of algebra yield

$$\frac{\Delta E}{E} = \sqrt{\frac{2}{3N}}.$$

In this equation, the factor of $2/3$ is specific to the system considered, an ideal gas. However, that $\Delta E/E \sim 1/\sqrt{N}$ is quite typical. We therefore see that for a macroscopic system, the fluctuations in energy are indeed very small: $\Delta E/E \sim 10^{-12}$. For all practical purposes, therefore, there is very little difference between a system interacting with a heat bath and one which is isolated.

This conclusion is one of the most useful results of statistical mechanics. Since the physical predictions will turn out to be identical, it is up to us to decide whether we will consider the system to be isolated or interacting with a heat reservoir. For most systems, it is easier to compute the partition function than the statistical weight. For this reason, it is usually more convenient to imagine the system to be interacting with a reservoir, instead of being isolated.

3.2.3   Application: N simple harmonic oscillators

We consider a system consisting of N identical SHO's interacting with a heat reservoir at a temperature T. To compute the system's thermodynamic quantities we must first evaluate the partition function,

$$Z = \sum_r e^{-\beta E_r}.$$

The first order of business is to figure out how the microstates of the entire system are specified. Since the quantum state of each oscillator is determined by a quantum number $n = 0, 1, 2, \ldots$, the quantum state of the whole system is determined by all the quantum numbers associated with all the oscillators. In other words,

$$r = \{n_1, n_2, \ldots, n_N\}.$$

The energy eigenvalue associated with this quantum state is then

$$E_r = (n_1 + \tfrac{1}{2})\hbar\omega + (n_2 + \tfrac{1}{2})\hbar\omega + \cdots + (n_N + \tfrac{1}{2})\hbar\omega.$$

How do we sum over all these quantum states? To do this we must sum over the set of all quantum numbers, that is, sum over all the possible values of each quantum number. To see this clearly, let us consider a simple system consisting of 2 SHO's. In this case the quantum states of the entire system are given by the set

$$\begin{aligned}
\{(n_1, n_2)\} \;=\;\; & \{(0,0), (0,1), \ldots, (0,\infty), \\
& (1,0), (1,1), \ldots, (1,\infty), \\
& \ldots \\
& (\infty,0), (\infty,1), \ldots, (\infty,\infty)\}.
\end{aligned}$$

To sum over all these states, we just proceed from left to right, and from top to bottom:

$$\text{sum} \quad = \quad \sum_{n_2=0}^{\infty} (0, n_2) + \sum_{n_2=0}^{\infty} (1, n_2) + \cdots + \sum_{n_2=0}^{\infty} (\infty, n_2)$$

$$= \quad \sum_{n_1=0}^{\infty} \left(\sum_{n_2=0}^{\infty} (n_1, n_2) \right).$$

In other words,

$$\sum_{r} = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} .$$

This is the answer for 2 oscillators; for N oscillators we have

$$\sum_{r} = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} .$$

The partition function, therefore, is given by

$$Z \quad = \quad \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} e^{-\beta[(n_1+\frac{1}{2})\hbar\omega + (n_2+\frac{1}{2})\hbar\omega + \cdots + (n_N+\frac{1}{2})\hbar\omega]}$$

$$= \quad \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} e^{-\beta\hbar\omega(n_1+\frac{1}{2})} e^{-\beta\hbar\omega(n_2+\frac{1}{2})} \cdots e^{-\beta\hbar\omega(n_N+\frac{1}{2})}.$$

Now, how do we evaluate this horrible multiple sum? To find out, let us once again consider our simple example, $N = 2$. Denoting $f(n) = \exp[-\beta\hbar\omega(n + \frac{1}{2})]$, we have

$$Z \quad = \quad \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} f(n_1) f(n_2)$$

$$= \quad \sum_{n_1=0}^{\infty} \Big[f(n_1) f(0) + f(n_1) f(1) + \cdots + f(n_1) f(\infty) \Big]$$

$$= \quad \sum_{n_1=0}^{\infty} f(n_1) \Big[f(0) + f(1) + \cdots + f(\infty) \Big]$$

$$= \quad \Big[f(0) + f(1) + \cdots + f(\infty) \Big] \sum_{n_1=0}^{\infty} f(n_1)$$

$$= \quad \Big[f(0) + f(1) + \cdots + f(\infty) \Big] \Big[f(0) + f(1) + \cdots + f(\infty) \Big]$$

$$= \quad \left(\sum_{n=0}^{\infty} f(n) \right)^2.$$

For N oscillators there are N sums, and this result generalizes to

$$Z \quad = \quad \left(\sum_{n_1=0}^{\infty} e^{-\beta\hbar\omega(n_1+\frac{1}{2})} \right) \left(\sum_{n_2=0}^{\infty} e^{-\beta\hbar\omega(n_2+\frac{1}{2})} \right) \cdots \left(\sum_{n_N=0}^{\infty} e^{-\beta\hbar\omega(n_N+\frac{1}{2})} \right)$$

$$= \quad \left(\sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+\frac{1}{2})} \right)^N ,$$

because the sums are all identical.

All that is left to do is to evaluate the single sum

$$\sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+\frac{1}{2})} = e^{-\frac{1}{2}\beta\hbar\omega} \sum_{n=0}^{\infty} \left(e^{-\beta\hbar\omega}\right)^n.$$

This is easy, because we can use the identity

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots = \sum_{n=0}^{\infty} x^n,$$

which is proved by just multiplying both sides by $1 - x$. We obtain

$$\sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+\frac{1}{2})} = \frac{e^{-\frac{1}{2}\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}.$$

Finally, collecting the results, we arrive at the partition function

$$Z = \frac{e^{-\frac{N}{2}\beta\hbar\omega}}{\left(1 - e^{-\beta\hbar\omega}\right)^N}. \tag{3.14}$$

From this we can compute every thermodynamic quantity of interest. This requires us to first take a log:

$$\ln Z = -\tfrac{1}{2}N\beta\hbar\omega - N\ln\left(1 - e^{-\beta\hbar\omega}\right).$$

Then, the thermodynamic energy is calculated as $E = -\partial\ln Z/\partial\beta$, or

$$E = N\hbar\omega\left(\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega} - 1}\right). \tag{3.15}$$

The entropy is then obtained as $S = k(\ln Z + \beta E)$, or

$$S = Nk\left[\frac{\beta\hbar\omega}{e^{\beta\hbar\omega} - 1} - \ln\left(1 - e^{-\beta\hbar\omega}\right)\right]. \tag{3.16}$$

Finally, the chemical potential is obtained from $\mu = -kT\partial\ln Z/\partial N$, and the result is

$$\mu = \tfrac{1}{2}\hbar\omega + kT\ln\left(1 - e^{-\beta\hbar\omega}\right). \tag{3.17}$$

It is instructive to look at these results in the limit of low temperatures ($kT \ll \hbar\omega$). For example,

$$E \simeq \tfrac{1}{2}N\hbar\omega + N\hbar\omega e^{-\hbar\omega/kT}.$$

The energy is therefore dominated by the zero-point energy ($\tfrac{1}{2}\hbar\omega$) of each oscillator. This makes sense, because at low temperatures, the vast majority of the oscillators are in the ground state $n = 0$. Similarly,

$$S = Nk\,\frac{\hbar\omega}{kT}\,e^{-\hbar\omega/kT}$$

and

$$\mu \simeq \tfrac{1}{2}\hbar\omega - kTe^{-\hbar\omega/kT}$$

at low temperatures. It is a good idea to carefully think about the physical meaning of these results; this will be left to the student.

In the limit of high temperatures ($kT \gg \hbar\omega$), our results reduce to

$$E \simeq NkT, \qquad S \simeq Nk\left[\ln\left(\frac{kT}{\hbar\omega}\right) + 1\right], \qquad \mu \simeq -kT\ln\left(\frac{kT}{\hbar\omega}\right).$$

Here also, the student is invited to think about the physical meaning of these results.

3.3   Gibbs distribution

The Gibbs distribution applies to a system interacting with a reservoir, when the interaction allows for an exchange of both heat and particles. Here it is convenient to introduce, apart from $\beta \equiv 1/kT$, the parameter

$$\boxed{\alpha = -\frac{\mu}{kT}}\,, \tag{3.18}$$

where μ is the chemical potential of the reservoir. The Gibbs distribution of probabilities is then given by

$$\boxed{p_r = \frac{1}{\mathcal{Z}}\, e^{-(\beta E_r + \alpha N_r)}, \qquad \mathcal{Z} = \sum_r e^{-(\beta E_r + \alpha N_r)}}\,. \tag{3.19}$$

The quantity $\mathcal{Z}$ is called the *grand partition function*, and it plays essentially the same role that the partition function plays in the Boltzmann distribution. In particular, we will see that all thermodynamic quantities can be computed directly from the grand partition function.

3.3.1   Thermodynamic quantities

We define the thermodynamic energy E exactly as before, as the system's mean energy. The only difference is that now, the average is computed with the Gibbs distribution. So

$$E \equiv \overline{E_r} = \sum_r p_r E_r = \frac{1}{\mathcal{Z}} \sum_r E_r e^{-(\beta E_r + \alpha N_r)}.$$

It is easy to show, as we have done before for the Boltzmann distribution, that E can be conveniently expressed as

$$\boxed{E = -\frac{\partial \ln \mathcal{Z}}{\partial \beta}}\,. \tag{3.20}$$

The first law of thermodynamics contains the term $\mu\, dN$, where N is the system's macroscopic number of particles. Since our system does not have a fixed number of particles, we must also define N as an averaged quantity. More precisely, we define N as

$$N = \overline{N_r} = \sum_r p_r N_r = \frac{1}{\mathcal{Z}} \sum_r N_r e^{-(\beta E_r + \alpha N_r)}.$$

The same steps that lead to Eq. (3.20) can be applied here, and we find

$$\boxed{N \equiv -\frac{\partial \ln \mathcal{Z}}{\partial \alpha}}\,. \tag{3.21}$$

We now calculate the entropy from the defining relation (3.4). Expressing $\ln p_r$ as $-(\ln \mathcal{Z} + \beta E_r + \alpha N_r)$, we obtain

$$S = k \sum_r p_r (\ln \mathcal{Z} + \beta E_r + \alpha N_r),$$

or

$$\boxed{S = k(\ln \mathcal{Z} + \beta E + \alpha N)}\,. \tag{3.22}$$

This implies that the Landau potential $\Omega = E - TS - \mu N$ is very closely related to the grand partition function. Indeed, a few steps of algebra yield

$$\boxed{\Omega(T, V, \mu) = -kT \ln \mathcal{Z}(T, V, \mu)} . \tag{3.23}$$

That Ω, and therefore also $\mathcal{Z}$, depend on the quantities T, V, and μ follows from the relation

$$d\Omega = -S\, dT - P\, dV - N\, d\mu,$$

derived in part A of this course (Sec. A7). This relation also gives definitions for S, P, and N in terms of partial derivatives of Ω:

$$\boxed{S = -\left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} \qquad P = -\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} \qquad N = \left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V}} . \tag{3.24}$$

Notice that the equations $S = -\partial\Omega/\partial T$ and $N = -\partial\Omega/\partial\mu$ are entirely equivalent to Eqs. (3.22) and (3.21), respectively.

So, if the grand partition function is known as a function of T, V, and μ, then the Landau potential is known, and *all other* thermodynamic quantities can be obtained by partial differentiation. In other words:

> For a system in thermo-chemical contact with a reservoir at temperature T and chemical potential μ, every thermodynamic quantity can be calculated directly from the grand partition function $\mathcal{Z}(T, V, \mu)$.

This is highly analogous to the situations encountered before.

3.3.2   Fluctuations

A system described by the Gibbs distribution interacts with a reservoir, and the interaction allows for an exchange of both energy and particles. We ask: how strongly can the mean values E and N deviate, at some instant of time, from the *actual* values for the energy and the number of particles? In other words, how large are the fluctuations? In the case of the Boltzmann distribution we saw that the fluctuations in energy are very small. Here we will show that the fluctuations are small also in the case of the Gibbs distribution.

We define ΔE and ΔN by

$$\boxed{(\Delta E)^2 = \overline{(E_r - E)^2}, \qquad (\Delta N)^2 = \overline{(N_r - N)^2}} . \tag{3.25}$$

We saw in Sec. 2 that in the case of the Boltzmann distribution,

$$\boxed{(\Delta E)^2 = -\left(\frac{\partial E}{\partial \beta}\right)_{\alpha,V}} . \tag{3.26}$$

This result is valid also for the Gibbs distribution. Instead of proving this explicitly, we will derive the analogous result for ΔN; the steps are virtually identical and will be left as an exercise for the student. We first express N as

$$N = \frac{\sum_r N_r e^{-(\beta E_r + \alpha N_r)}}{\sum_r e^{-(\beta E_r + \alpha N_r)}},$$

and compute its partial derivative with respect to α:

$$-\frac{\partial N}{\partial \alpha} = \frac{\sum_r N_r{}^2 e^{-(\beta E_r + \alpha N_r)}}{\sum_r e^{-(\beta E_r + \alpha N_r)}} - \left(\frac{\sum_r N_r e^{-(\beta E_r + \alpha N_r)}}{\sum_r e^{-(\beta E_r + \alpha N_r)}}\right)^2 .$$

We may now identify the first term of the right-hand side with $\overline{N_r{}^2}$, and the second term with $\overline{N_r}^2$. Since

$$\overline{(N_r - N)^2} = \overline{N_r{}^2} - \overline{N_r}^2,$$

we have established that

$$\boxed{(\Delta N)^2 = -\left(\frac{\partial N}{\partial \alpha}\right)_{\beta,V}}. \tag{3.27}$$

To demonstrate that the fluctuations are typically very small, we evaluate Eq. (3.27) for the specific case of an ideal gas. We must first express N in terms of α and β; otherwise, we might get confused when taking the partial derivative. We recall that for an ideal gas,

$$\alpha = -\frac{\mu}{kT} = \ln\left[\frac{V}{(2\pi\hbar)^3 N}\left(\frac{2\pi m}{\beta}\right)^{3/2}\right],$$

which implies

$$N = \frac{V}{(2\pi\hbar)^3}\left(\frac{2\pi m}{\beta}\right)^{3/2} e^{-\alpha}.$$

This immediately gives us $\partial N/\partial \alpha = -N$, and

$$\frac{\Delta N}{N} = \frac{1}{\sqrt{N}}.$$

We see that the fluctuation in the number of particles is indeed very small.

We shall now work out an expression for the energy fluctuation. To evaluate Eq. (3.26) we express $E = \frac{3}{2}NkT$ as

$$E = \frac{3}{2}\frac{V}{(2\pi\hbar)^3}\frac{(2\pi m)^{3/2}}{\beta^{5/2}} e^{-\alpha},$$

which implies that $\partial E/\partial \beta = -\frac{5}{2}E/\beta$. Finally, we arrive at

$$\frac{\Delta E}{E} = \sqrt{\frac{5}{3N}}.$$

This is slightly different from the result of Sec. 2, $\Delta E/E = \sqrt{2/3N}$. The difference is due to the fact that the fluctuations in the number of particles induce additional fluctuations in the energy, so that $(\Delta E)_{\text{Gibbs}} > (\Delta E)_{\text{Boltzmann}}$.

We conclude that for all practical purposes, the physical predictions obtained using the Gibbs distribution are virtually identical to those obtained using the Boltzmann distribution. Even though the number of particles is subject to fluctuations, these are so small that they can be safely ignored.

3.4   Classical statistics

We now carry on with our formulation of classical statistical mechanics, which we began in Sec. B4. We shall consider two topics. First, we will describe how to compute the partition function in the classical limit. Second, we will formulate and prove the very powerful *equipartition theorem* of classical statistics.

3.4.1   Partition function

Our goal here is to find an expression for the partition function which applies in the classical limit. We start with Eq. (3.11),

$$Z = \int e^{-\beta \mathcal{E}} \omega(\mathcal{E}) \, d\mathcal{E},$$

where $\omega(\mathcal{E}) \, d\mathcal{E} = \Omega(\mathcal{E}; d\mathcal{E})$ is the number of microstates accessible to the classical system if its energy is in the interval between $\mathcal{E}$ and $\mathcal{E} + d\mathcal{E}$. An expression for this was found in Sec. B4:

$$\omega(\mathcal{E}) \, d\mathcal{E} = \frac{1}{(2\pi\hbar)^f} \int_R dp_1 \cdots dp_f dq_1 \cdots dq_f,$$

where f denotes the number of degrees of freedom. (This expression for the density of states does not account for the possible indistinguishability of the particles. If the system is composed of indistinguishable particles, then this ω must be divided by $N!$, where N is the number of particles.) In this expression, the domain R of integration is determined by the fact that $H(p_1, \ldots, q_f)$, the system's Hamiltonian, must be bounded by $\mathcal{E} < H < \mathcal{E} + d\mathcal{E}$. Since the computation of Z involves integrating over all values of the energy, it is clear that Z can be written in terms of an integral extending over *all* of phase space. As a matter of fact, combining the last two equations gives

$$\boxed{Z = \frac{1}{(2\pi\hbar)^f} \int e^{-\beta H(p_1, \ldots, q_f)} \, dp_1 \cdots dp_f dq_1 \cdots dq_f}, \qquad (3.28)$$

where, as was just stated, the domain of integration extends over the entire phase space of the classical system. Equation (3.28) is the desired expression for the classical partition function. (Again, if the system is composed of indistinguishable particles, then this Z must be divided by $N!$, where N is the number of particles.)

3.4.2   The ideal gas — again

As an example of how to use Eq. (3.28), we now compute the classical partition function for an ideal gas. As we shall see, this calculation is far less laborious than the calculation of the statistical weight (no need for the "volume" of a $3N$-dimensional spherical shell!).

   The Hamiltonian for an ideal gas is

$$H = \frac{1}{2m} \left(p_1{}^2 + \cdots + p_{3N}{}^2 \right).$$

Since H does not depend on the coordinates, the integrations over the dq's can be factored out, and they give

$$\int dq_1 \cdots dq_{3N} = \left(\int dq_1 dq_2 dq_3 \right)^N = V^N.$$

Now, because the Boltzmann factor factorizes into

$$\exp\left(-\frac{\beta p_1{}^2}{2m} \right) \cdots \exp\left(-\frac{\beta p_{3N}{}^2}{2m} \right),$$

the momentum integrations evaluate to

$$\int e^{-\beta H} \, dp_1 \cdots dp_{3N} = \left(\int_{-\infty}^{\infty} e^{-\beta p^2 / 2m} \, dp \right)^{3N} = \left(\frac{2\pi m}{\beta} \right)^{3N/2},$$

where we have used $\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}$. Finally, gathering the results yields

$$Z = \frac{1}{N!} \left[\frac{V}{(2\pi\hbar)^3} \left(\frac{2\pi m}{\beta} \right)^{3/2} \right]^N, \qquad (3.29)$$

where the factor $1/N!$ was inserted to account for the indistinguishability of the particles. Using the Stirling approximation, we also obtain

$$\ln Z = N \ln \left[\frac{V}{(2\pi\hbar)^3 N} \left(\frac{2\pi m}{\beta} \right)^{3/2} \right] + N.$$

The thermodynamic relations for an ideal gas can be derived, in the usual way, from the partition function. The mean energy is given by $E = -\partial \ln Z / \partial \beta$. Since $\ln Z = -\frac{3}{2} N \ln \beta$ plus terms not involving β, this gives

$$E = \frac{3}{2} NkT.$$

The entropy is given by $S = k(\ln Z + \beta E)$, which evaluates to

$$S = Nk \ln \left[\frac{V}{(2\pi\hbar)^3 N} (2\pi mkT)^{3/2} \right] + \tfrac{5}{2} Nk.$$

The pressure is obtained from $P = kT \partial \ln Z / \partial V$. Since $\ln Z = N \ln V$ plus terms not involving V, this gives

$$PV = NkT.$$

And finally, the chemical potential is given by $\mu = -kT \partial \ln Z / \partial N$, which implies

$$\mu = -kT \ln \left[\frac{V}{(2\pi\hbar)^3 N} (2\pi mkT)^{3/2} \right].$$

These results are all idential to the ones obtained in part B of this course, on the basis of the statistical weight. However, the calculations presented here were much more straightforward.

3.4.3   Equipartition theorem

As we have seen repeatedly, the macroscopic energy of an ideal gas is given, in the classical limit, by $E = \frac{3}{2}NkT$. On the other hand, in Sec. 2c above we obtained the expression $E = NkT$ for the energy of a system of simple harmonic oscillators (also in the classical limit). These results suggest that in the classical limit, the total energy of a system is always distributed in such a way that the mean energy per particle is of order kT. The precise version of this statement is known as the *equipartition theorem*, which we now formulate and prove.

Suppose that the Hamiltonian of a classical system can be written as

$$H(p_1, \ldots, p_f, q_1, \ldots, q_f) = h(\xi) + H'(p_1, \ldots, p_f, q_1, \ldots, q_f),$$

where h depends only on ξ, while H' does not at all depend on ξ. Here, ξ is *any one* of the phase-space coordinates; it can equally well be any one of the p's or any one of the q's. Suppose also that h is *quadratic* in ξ:

$$h(\xi) = a\xi^2,$$

where a is a constant. Then the equipartition theorem states that the mean value of h, with respect to the Boltzmann distribution, is always given by

$$\boxed{\bar{h} = \tfrac{1}{2}kT} \,, \tag{3.30}$$

irrespective of the form of H'.

The equipartition theorem implies that each quadratic term in the Hamiltonian contributes $\tfrac{1}{2}kT$ to the mean total energy. Because the Hamiltonian of an ideal gas contains quadratic terms *only*, and a number $3N$ of them, the theorem implies that $E = \tfrac{3}{2}NkT$. Similarly, the Hamiltonian of a system of simple harmonic oscillators contains $2N$ quadratic terms, which gives $E = NkT$. It also follows from the equipartition theorem that the mean *kinetic* energy of a classical system is always equal to $\tfrac{1}{2}fkT$, where f is the number of degrees of freedom.

We now proceed with the proof of the theorem. By definition, the mean value of h is given by

$$\bar{h} = \frac{\int e^{-\beta H} h \, dp_1 \cdots dq_f}{\int e^{-\beta H} \, dp_1 \cdots dq_f}.$$

If we express H as $h + H'$, this becomes

$$\bar{h} = \frac{\int e^{-\beta h} h \, d\xi \int e^{-\beta H'} \, dp_1 \cdots dq_f}{\int e^{-\beta h} \, d\xi \int e^{-\beta H'} \, dp_1 \cdots dq_f},$$

where it is understood that $d\xi$ is excluded from $dp_1 \cdots dq_f$. The integrals not involving ξ cancel out, and we obtain

$$\bar{h} = \frac{\int e^{-\beta h} h \, d\xi}{\int e^{-\beta h} \, d\xi} = -\frac{\partial}{\partial \beta} \ln \int e^{-\beta h} \, d\xi.$$

We now use the fact that $h = a\xi^2$, and define a new integration variable $u = \sqrt{\beta}\xi$. A few steps of algebra then give

$$\bar{h} = -\frac{\partial}{\partial \beta} \left(-\tfrac{1}{2} \ln \beta + \ln \int e^{-au^2} \, du \right) = \frac{1}{2\beta},$$

which proves the theorem. Notice that we did not have to evaluate a single integral in order to establish this beautiful result!

3.5   The meaning of probability

We now have at our disposal the fundamental tools of statistical mechanics, and the rest of the course will focus on various applications. In this section we examine the following conceptual question: What does it mean to say that a thermodynamic system has a probability p_r of being found in a microstate ψ_r? We shall give two answers, and then argue that these two answers must be equivalent.

For the first answer, we suppose that the microstate of our system is measured at many different times t_1, t_2, ..., t_n, where n is very large. (We assume that the time intervals are sufficiently long that the system has time to return to equilibrium between successive measurements.) Then we can define the probability of finding the system in the microstate ψ_r as

$$p_r = \frac{\text{number of times the system is measured to be in } \psi_r}{\text{total number of measurements}}.$$

This definition is obviously very close to what would be an *experimental* determination of the probabilities. Averages constructed with these probabilities are called "time averages", for obvious reasons.

The second answer is quite different. Here we imagine that we have in our possession n identical copies of our thermodynamic system, and that we measure the microstate of each of these at a single moment of time. Because these systems are all identical, they are all in the same macrostate. However, there is no reason to assume that the systems are all in the same microstate. Instead, we may define the probability of finding our original system in the microstate ψ_r as

$$p_r = \frac{\text{number of copies found in microstate } \psi_r}{\text{total number of copies}}.$$

This obviously is a less practical definition, but it has proven to be an extremely fruitful one for the development of the *theory* of statistical mechanics. It is usual to refer to the n copies of the original system as an *ensemble* of systems, and averages constructed in this way are called "ensemble averages".

It is intuitively clear that those two ways of defining probabilities are equivalent: making measurements at many different times on the same system is pretty much the same thing as making simultaneous measurements on many identical systems. The hypothesis according to which this is true is known as the *ergotic hypothesis*, and again, the validity of this hypothesis is intuitively clear. However, the proof of its validity could never be obtained within the framework of statistical mechanics. In fact, counterexamples can be formulated in certain special situations. Nevertheless, the validity of the ergotic hypothesis is widely accepted, and it can be introduced as an additional postulate of statistical mechanics.

In the remainder of this course it will be helpful to use the language of ensembles. We have so far encountered three types of ensembles. The *microcanonical ensemble* designates a large number of identical, *isolated* systems; in the microcanonical ensemble, probabilities are given by

$$p_r = \frac{1}{\Omega},$$

where Ω is the statistical weight. The *canonical ensemble* designates a large number of identical systems in *thermal interaction* (V and N fixed) with a reservoir; in the canonical ensemble, probabilities are given by the Boltzmann distribution,

$$p_r = \frac{1}{Z} \exp\bigl(-E_r/kT\bigr).$$

Finally, the *grand canonical ensemble* designates a large number of identical systems in *thermo-chemical interaction* (only V fixed) with a reservoir; in the grand canonical ensemble, probabilities are given by the Gibbs distribution,

$$p_r = \frac{1}{\mathcal{Z}} \exp\Bigl[-\bigl(E_r - \mu N_r\bigr)/kT\Bigr].$$

3.6 Brief summary of statistical mechanics

We use the following symbols: $\alpha \equiv -\mu/kT$, $\beta \equiv 1/kT$. The index r labels the microstates accessible to the system; these have energy eigenvalues E_r and particle numbers N_r. The sum $\sum_r$ is over all these microstates.

3.6.1 Boltzmann distribution

Valid for systems in thermal interaction (V and N fixed) with a reservoir.

- Boltzmann distribution:
$$p_r = \frac{1}{Z} e^{-\beta E_r}$$

- Partition function:

$$Z(\beta, V, N) = \sum_r e^{-\beta E_r} = \int \omega(\mathcal{E}) e^{-\beta \mathcal{E}} \, d\mathcal{E}$$

- Helmholtz free energy:
$$F = -kT \ln Z$$

- Mean energy:
$$E = -\frac{\partial \ln Z}{\partial \beta}$$

- Entropy:
$$S = k(\ln Z + \beta E)$$

- Pressure:
$$P = kT \frac{\partial \ln Z}{\partial V}$$

- Chemical potential:
$$\mu = -kT \frac{\partial \ln Z}{\partial N}$$

- Energy fluctuation:
$$(\Delta E)^2 = -\frac{\partial E}{\partial \beta}$$

3.6.2   Gibbs distribution

Valid for systems in thermo-chemical interaction (only V fixed) with a reservoir.

- Gibbs distribution:
$$p_r = \frac{1}{\mathcal{Z}} e^{-(\beta E_r + \alpha N_r)}$$

- Grand partition function:

$$\mathcal{Z}(\alpha, \beta, V) = \sum_r e^{-(\beta E_r + \alpha N_r)}$$

- Landau potential:
$$\Omega = -kT \ln \mathcal{Z}$$

- Mean number of particles:
$$N = -\frac{\partial \ln \mathcal{Z}}{\partial \alpha}$$

- Mean energy:
$$E = -\frac{\partial \ln \mathcal{Z}}{\partial \beta}$$

- Entropy:
$$S = k(\ln \mathcal{Z} + \beta E + \alpha N)$$

- Pressure:
$$P = kT \frac{\partial \ln \mathcal{Z}}{\partial V}$$

- Fluctuations:
$$(\Delta N)^2 = -\frac{\partial N}{\partial \alpha}, \qquad (\Delta E)^2 = -\frac{\partial E}{\partial \beta}$$

3.7   Problems

1. We continue our study of a system consisting of N spins, each capable of having the values $\pm\varepsilon$ for its energy. We now imagine that the system is in contact with a heat reservoir at temperature T.

   The state of the nth spin is specified by the quantum number $m_n = \pm 1$, so that its energy is $m_n\varepsilon$. Correspondingly, the microstate of the entire system is specified by the set $\{m_n\}$ of all N quantum numbers. We denote this set by r. If the system is in a specified microstate r, then its total energy is

$$E_r = \sum_{n=1}^{N} m_n\varepsilon.$$

   a) Show that the partition function is given by

$$Z = \left[2\cosh(\beta\varepsilon)\right]^N,$$

   where $\beta = 1/kT$.

   b) Calculate the mean energy E of the system. How does your answer compare with your result from last week?

   c) Calculate ΔE, the energy fluctuation. Compare ΔE with E itself. Do the fluctuations ever become large?

2. (Reif 7.2) An ideal monatomic gas of N particles, each of mass m, is in thermal equilibrium at absolute temperature T. The gas is contained in a cubical box of side L, whose top and bottom are parallel to the earth's surface. The effect of the earth's uniform gravitational field on the particles should be considered. You may assume that the gravitational potential is zero at the bottom of the box, so that a particle at a distance y from the bottom has potential energy mgy, where g is the gravitational acceleration.

   a) Calculate the partition function for this gas, starting from the classical result

$$Z = \frac{1}{N!(2\pi\hbar)^{3N}} \int e^{-\beta H} d\vec{p}_1 \cdots d\vec{p}_N \, d\vec{r}_1 \cdots d\vec{r}_N,$$

   where H is the Hamiltonian, and where $\vec{p}_1$ and $\vec{r}_1$ are respectively the momentum vector and position vector of the first particle; $d\vec{p}_1$ stands for $dp_{1x}dp_{1y}dp_{1z}$.

   b) Obtain the equation of state for this gas (pressure as a function of L and T). Check that your answer reduces to the familiar result $PV = NkT$ in the case of zero gravity.

3. The *virial theorem* of classical mechanics states that if a body moves under the influence of a central potential $V(r) \propto r^n$, where r is the distance from the centre to the body, then $\bar{K}$, its mean kinetic energy, and $\bar{V}$, its mean potential energy, are related by

$$\bar{K} = \frac{n}{2}\,\bar{V}.$$

   Consider a system consisting of N moving bodies, all under the influence of this potential, in thermal equilibrium at a temperature T.

   a) What is $\bar{E}$, the system's mean total energy? (Express your result in terms of N and T.)

b) What is $\bar{E}$ for the specific case of a gravitational interaction? Why is $\bar{E}$ negative?

4. In this problem we consider a system consisting of N simple harmonic oscillators in thermal equilibrium at a temperature T. We assume that the oscillators are all identical (so that they all have the same frequency ω), distinguishable, and one dimensional (so that they each possess a single degree of freedom).

   a) How high must the temperature be for the system to behave classically? (Give a precise criterion.)

   b) In the classical limit, calculate the partition function for this system.

   c) In the classical limit, calculate E, the system's mean energy.

   d) In the classical limit, calculate S, the system's entropy.

   e) In the classical limit, calculate μ, the system's chemical potential.

5. A *rotor* is a quantum mechanical system with Hamiltonian

$$H = \frac{\boldsymbol{J}^2}{2I},$$

where $\boldsymbol{J}$ is the angular momentum operator, and I a constant (the rotor's moment of inertia). The rotor's eigenstates are $|j, m\rangle$, such that

$$\boldsymbol{J}^2 |j, m\rangle = j(j+1)\hbar^2 |j, m\rangle, \qquad J_z |j, m\rangle = m\hbar |j, m\rangle,$$

where $j = 0, 1, 2, \ldots$, and $m = -j, -j+1, \ldots, j-1, j$.

Calculate the partition function for a system consisting of a single rotor in thermal equilibrium with a heat bath at a temperature T. You may assume that T is sufficiently high that the discrete sum over j can be turned into an integral.

CHAPTER 4
INFORMATION THEORY

Information theory is a reformulation of statistical mechanics in terms of the *missing information* associated with a thermodynamic system whose microstate is not known. This alternative approach provides a completely different derivation of the Boltzmann and Gibbs distributions. Although it is much more formal than what was presented in parts B and C, this formulation of statistical mechanics sheds new light on the fundamental meaning of the entropy.

4.1 Missing information

How do we quantify the amount of missing information associated with a system whose microstate is not known? We answer this question in two steps.

4.1.1 Uniform probabilities

Consider a *problem of choice*: n possibilities are given, of which only one is true; which one it is is not known. (For example, a thermodynamic system can be in any one of its n accessible microstates.)

What is I, the missing information, if the n possibilities are all equally likely? (For example, the thermodynamic system could be isolated.) To answer this, let us be guided by a number of properties that the missing information should possess. First, I should depend on n, and I should increase as the number of possibilities is increased:

$$I(n_2) > I(n_1) \text{ if } n_2 > n_1.$$

Second, when the number of possibilities is equal to one, there should be no missing information:

$$I(1) = 0.$$

The third property requires more discussion. Consider two *independent* problems of choice, one with n_1 possibilities, the other with n_2 possibilities. (For example, a first thermodynamic system has n_1 accessible states, while a second system, not interacting with the first, has n_2 accessible states.) Clearly, the total number of possibilities for the combined problem of choice is $n_1 n_2$. The third property of the missing information is that it should be *additive*:

$$I(n_1 n_2) = I(n_1) + I(n_2).$$

There is a *unique* choice of function $I(n)$ that satisfies all three properties. It is

$$\boxed{I(n) = k \ln(n)}, \tag{4.1}$$

where k, for now, is an arbitrary constant.

4.1.2 Assigned probabilities

We now generalize the discussion to the case of unequal probabilities. In our problem of choice, we now label the various possibilities with the index r, and we assign a probability p_r to each possibility. (For example, a system has a probability p_r of being found in the quantum state ψ_r.) We assume *nothing* regarding the probability distribution, except that it must be properly normalized,

$$\sum_r p_r = 1.$$

In particular, we do *not* assume that the probabilities are given by the Boltzmann or Gibbs distributions.

    The meaning of these probabilities is taken to be the following. Suppose that we have a large number N of identical problems of choice — an *ensemble* of problems. (For example, an ensemble of thermodynamic systems consists of N identical copies of the original system.) Then the probabilities are defined so that when $N \to \infty$,

$$p_r = \frac{N_r}{N},$$

where N_r is the number of times that possibility r turns out to be true. So if the probability distribution is known, then we know that among the N problems of choice, possibility r will be true in N_r cases.

    However, what we do *not* know is *which* of the N problems are the ones for which possibility r is true. This represents missing information, and to evaluate $I(m)$, we need to first calculate m, the number of ways our ensemble of problems can be prepared so that possibility r will be true for N_r of them. A few moments of deep thought will convince one that m is equal to the number of ways in which $N_1 + N_2 + \cdots = N$ objects, of which N_1 are indistinguishably of type 1, and of which N_2 are indistinguishably of type 2, $\ldots$, can be allocated to N slots.

    If we were dealing with distinguishable objects, m would be equal to

$$N \times (N-1) \times \cdots \times 1 = N!,$$

since N is the number of ways the first object can be chosen among the N objects to fill the first slot, while $N-1$ is the number of ways the second object can be chosen among the remaining $N-1$ objects to fill the second slot, and so on. In other words, $m = N!$ for distinguishable objects. But since the N_1 objects of type 1 are indistinguishable, we must divide this by $N_1!$, the number of permutations of these objects among themselves. And since the N_2 objects of type 2 are also indistinguishable, we must also divide this by $N_2!$, the number of permutations of these objects among themselves. Continuing in this vein, we arrive at

$$m = \frac{N!}{N_1! N_2! \cdots}.$$

    The missing information is therefore

$$
\begin{aligned}
I(m) &= k \ln\left(\frac{N!}{N_1! N_2! \cdots}\right) \\
&= k\left[\ln(N!) - \ln(N_1!) - \ln(N_2!) - \cdots\right] \\
&= k\left[\ln(N!) - \sum_r \ln(N_r!)\right].
\end{aligned}
$$

Since we eventually take the limit $N \to \infty$ (which implies $N_r \to \infty$), we may use the Stirling approximation $k \ln(N!) = N \ln N - N$. This gives

$$I(m) = k\left(N \ln N - \sum_r N_r \ln N_r\right),$$

where the relation $N = \sum_r N_r$ was used. We now write $N_r = Np_r$, and after simplification, $I(m)$ becomes

$$I(m) = -Nk \sum_r p_r \ln p_r.$$

This is the missing information associated with our ensemble of N identical problems of choice. The missing information *per problem* is then

$$\boxed{I = -k \sum_r p_r \ln p_r} \ . \tag{4.2}$$

This result holds for arbitrarily assigned probabilities. It also holds for uniform probabilities (a special case), in which case Eq. (4.2) reduces to Eq. (4.1), as a few steps of algebra confirm.

4.2   Entropy

The discussion of the previous section was quite general: it did not matter what the problem of choice was, and we were able to derive an explicit expression for the missing information. Now we consider the specific application in which the problem of choice is a thermodynamic system. The various possibilities are then the various microstates accessible to the system. We assume that probabilities can be assigned to these microstates, but we do not yet assume that these are given by the Boltzmann or Gibbs distributions. In fact, we will *derive* these distributions below.

    The first postulate of the information-theory formulation of statistical mechanics is that the thermodynamic entropy *is precisely* the missing information associated with the system, whose microstate in not known. Mathematically, therefore,

$$\boxed{S = I = -k \sum_r p_r \ln p_r} \ , \tag{4.3}$$

where k is now defined to be Boltzmann's constant.

    This identification of the entropy with the missing information is quite compelling. We have seen that the microscopic description of a thermodynamic system typically involves the specification of $3N$ quantum numbers, one for each degrees of freedom. We have also seen that the macroscopic description of the same system involves the specification of only a very small number of quantities. (For example: energy, volume, and number of particles.) The reduction from a microscopic description, which would be utterly impractical, to a macroscopic one, which is much more useful, necessarily implies a loss of (a lot of) information. This lost information is precisely what the entropy is. This truly is a beautiful realization, which endows the entropy with a very concrete physical meaning.

4.3   Boltzmann and Gibbs distributions

We have encountered Eq. (4.3) before, in part C of this course. There is an important difference, however. It is that in part C, the probabilities were specifically given by the Boltzmann and Gibbs distributions, whereas no such assumption was made here. In this section we use Eq. (4.3) to *derive* the Boltzmann and Gibbs distributions.

4.3.1   Maximum missing information

The second postulate of the information-theory formulation of statistical mechanics
is that a system *in equilibrium* has a probability distribution which is such that the
missing information is *maximized*. The distribution is further constrained by the
following relations:

$$\sum_r p_r = 1, \qquad \sum_r E_r p_r = E, \qquad \sum_r N_r p_r = N.$$

The mean energy E and the mean number of particles N are therefore assumed
to be known. This represents the *only* information available about the system.
Maximizing the missing information therefore amounts to assuming that we really
do not have any other information about the system.

The probability distribution is therefore determined by the condition

$$\boxed{\delta I = -k \sum_r (\ln p_r + 1)\, \delta p_r = 0}\,, \tag{4.4}$$

where the variations δp_r are subject to the constraints

$$\boxed{\sum_r \delta p_r = 0, \qquad \sum_r E_r\, \delta p_r = 0, \qquad \sum_r N_r\, \delta p_r = 0}\,. \tag{4.5}$$

These constraints imply that the δp_r's are not all independent. (If they were,
we would have to conclude that $\ln p_r = -1$, which clearly is impossible.) The
maximization must be carried out using the method of Lagrange multipliers.

4.3.2   Lagrange multipliers

Before we return to the problem of maximizing I subject to the constraints (4.5),
we first describe how to use the method of Lagrange multipliers.

Suppose we want to extremize the function $f(x, y, z)$ when the coordinates x,
y, and z are subject to the constraint $g(x, y, z) = 0$. (Geometrically, the equation
$g = 0$ describes a two-dimensional surface in three-dimensional space.) If f is to be
an extremum, then

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz = 0.$$

If the displacements were all independent, then we would conclude that all the par-
tial derivatives must be zero: $\partial f/\partial x = \partial f/\partial y = \partial f/\partial z = 0$. But the displacements
are not independent, because $g = 0$ implies

$$dg = \frac{\partial g}{\partial x}\, dx + \frac{\partial g}{\partial y}\, dy + \frac{\partial g}{\partial z}\, dz = 0.$$

(The displacements must lie on the surface $g = 0$.) This means that dz can be
expressed in terms of dx and dy, which are the truly independent displacements.

Consider the linear combination $df + \lambda dg$, where λ is some number (the Lagrange
multiplier). This combination must obviously vanish, because $df = dg = 0$. So

$$\left(\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y}\right) dy + \left(\frac{\partial f}{\partial z} + \lambda \frac{\partial g}{\partial z}\right) dz = 0.$$

The clever trick associated with the method of Lagrange multipliers is that we can
choose λ in such a way that

$$\frac{\partial f}{\partial z} + \lambda \frac{\partial g}{\partial z} = 0.$$

Having made this choice, the preceding equation becomes

$$\left(\frac{\partial f}{\partial x} + \lambda\frac{\partial g}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \lambda\frac{\partial g}{\partial y}\right) dy = 0.$$

But since the displacements dx and dy are independent, we must have

$$\frac{\partial f}{\partial x} + \lambda\frac{\partial g}{\partial x} = 0,$$
$$\frac{\partial f}{\partial y} + \lambda\frac{\partial g}{\partial y} = 0.$$

Maximization has been achieved. Notice that all three equations are of the same form, even though the first defines λ, while the second and third take care of the maximization.

The method therefore gives a very convenient way of maximizing a function f when the displacements are subject to the constraint $g = 0$: By taking the linear combination $df + \lambda dg$, one can pretend that the displacements are all independent, and set all coefficients to zero in the usual way. The value of the Lagrange multiplier is then determined from any one of the resulting equations.

4.3.3   Probability distributions

Going back to Eqs. (4.4) and (4.5), we see that in our variation of I, the displacements δp_r are subject to three constraints. This means that we must introduce three Lagrange multipliers, which we denote $\lambda - 1$, α, and β, respectively. The method instructs us to take the linear combination

$$\sum_r (\ln p_r + 1)\, \delta p_r + (\lambda - 1)\sum_r \delta p_r + \beta\sum_r E_r\, \delta p_r + \alpha\sum_r N_r\, \delta p_r = 0,$$

or

$$\sum_r \left(\ln p_r + \lambda + \beta E_r + \alpha N_r\right)\delta p_r = 0.$$

We can now pretend that the δp_r's are all independent, which means that every term in the sum must vanish separately. Thus, $\ln p_r = -(\lambda + \beta E_r + \alpha N_r)$, or

$$\boxed{p_r = e^{-\lambda}e^{-(\beta E_r + \alpha N_r)}} \ . \tag{4.6}$$

This is the Gibbs distribution! The Gibbs distribution reduces to the Boltzmann distribution when N_r is the same for all quantum states, that is, if the system has a fixed number of particles.

4.3.4   Evaluation of the Lagrange multipliers

We still have to evaluate the Lagrange multipliers. The first one, λ, is determined by making sure that the probabilities are properly normalized. This gives

$$\boxed{e^\lambda = \sum_r e^{-(\beta E_r + \alpha N_r)} = \mathcal{Z}} \ , \tag{4.7}$$

where $\mathcal{Z}$ is the grand partition function. The second multiplier, β, is determined by making sure that the mean energy is properly given by the known value E. This gives

$$\boxed{E = \frac{1}{\mathcal{Z}}\sum_r E_r e^{-(\beta E_r + \alpha N_r)} = -\frac{\partial \ln \mathcal{Z}}{\partial \beta}} \ . \tag{4.8}$$

Finally, α is determined by making sure that the mean number of particles is properly given by the known value N:

$$N = \frac{1}{\mathcal{Z}} \sum_r N_r e^{-(\beta E_r + \alpha N_r)} = -\frac{\partial \ln \mathcal{Z}}{\partial \alpha} . \qquad (4.9)$$

These are all the usual relations, which were derived in a completely different way in part C.

4.3.5   The first law

It is quite clear that the β and α multipliers are related to the system's temperature and chemical potential. In fact, we know the precise relations from part C. It is, however, not at all difficult to derive these relations from scratch, on the basis of what we have learned in this section.

We start from Eq. (4.8). Recalling that the energy eigenvalues may depend on the system's volume, this equation implies

$$dE = -P \, dV + \sum_r E_r \, dp_r ,$$

where P, for now, is merely short-hand for the quantity

$$-\sum_r \frac{\partial E_r}{\partial V} \, p_r .$$

Similarly, Eq. (4.9) implies

$$dN = \sum_r N_r \, dp_r .$$

Next, we go back to Eq. (4.3), the definition of the entropy. Differentiation yields

$$dS = -k \sum_r (\ln p_r + 1) \, dp_r = -k \sum_r \ln p_r \, dp_r ,$$

because $\sum_r dp_r = 0$. Substituting Eq. (4.6), this becomes

$$dS = k\beta \sum_r E_r \, dp_r + k\alpha \sum_r N_r \, dp_r .$$

Finally, combining the results, we arrive at

$$T \, dS = kT\beta(dE + P \, dV) + kT\alpha \, dN .$$

This has exactly the same form as the first law, $T \, dS = dE + P \, dV - \mu \, dN$, provided that

$$\beta = \frac{1}{kT} , \qquad \alpha = -\frac{\mu}{kT} . \qquad (4.10)$$

These are the usual relations. We also see that the quantity P introduced above is really the thermodynamic pressure.

4.4   Conclusion

We see that all the standard results of statistical mechanics can be derived from an information-theory approach, which is based upon two fundamental postulates. The first postulate states that the entropy of a thermodynamic system is precisely

equal to the missing information associated with the fact the system's microstate is not known. The second postulate states that the probability distribution that characterizes a system in equilibrium is the one which maximizes the missing information. (The maximization is subject to the constraint that the mean energy and the mean number of particles must be kept fixed during the variation.) This probability distribution turns out to be the Gibbs distribution, or its special case, the Boltzmann distribution.

Information theory provides a remarkably elegant, but formal, formulation of statistical mechanics. But more important, it highlights the direct physical meaning of the entropy.

CHAPTER 5
PARAMAGNETISM

We now have at hand the basic tools of statistical mechanics, and it is time to apply these tools to the thermodynamic description of various physical systems. We begin with a rather interesting application: magnetism, or more precisely, paramagnetism.

5.1 Magnetism

Magnetism in matter is caused by electric currents generated by moving electrons inside atoms and molecules. We shall now review the basic notions.

5.1.1 Magnetic moment

Any loop of current produces a magnetic field. Far from the loop, the magnetic vector potential $\boldsymbol{A}$ is given by

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \frac{\boldsymbol{\mu} \times \boldsymbol{r}}{r^3},$$

where $\boldsymbol{r}$ is the position at which the magnetic field $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$ is measured, and $r = |\boldsymbol{r}|$. The vectorial quantity $\boldsymbol{\mu}$ is called the *magnetic moment* of the loop, and it is given by

$$\boldsymbol{\mu} = (\text{current})(\text{area inside loop})\,\boldsymbol{n},$$

where $\boldsymbol{n}$ is a unit vector pointing in the direction normal to the loop, in accordance with the right-hand rule. Thus, the magnetic moment is the fundamental quantity which characterizes the strength of the magnetic field.

5.1.2 Potential energy

We consider a loop with magnetic moment $\boldsymbol{\mu}$ immersed in a uniform external magnetic field $\boldsymbol{B}_{\text{ext}}$. We assume that the external field is much stronger than the magnetic field produced by the loop itself.

In such a situation, the loop experiences a torque which tends to align $\boldsymbol{\mu}$ in the direction of $\boldsymbol{B}_{\text{ext}}$. (This means that the loop tends to align itself perpendicular to the field lines.) This torque is given by $\boldsymbol{\tau} = \boldsymbol{\mu} \times \boldsymbol{B}_{\text{ext}}$.

To rotate the loop in the external field therefore requires work, which increases the magnetic potential energy of the loop. This potential energy can be shown to be given by

$$U = -\boldsymbol{\mu} \cdot \boldsymbol{B}_{\text{ext}}.$$

The potential energy is minimum when $\boldsymbol{\mu}$ points in the same direction as the applied field. It is maximum when $\boldsymbol{\mu}$ points in the opposite direction.

5.1.3   Magnetic moments of atoms and molecules

Most atoms and molecules possess a magnetic moment, which we now calculate. Let us start with a simple *classical* model: a single electron moves on a circular orbit around a nucleus. The orbit represents a loop of current. The current is $-e/P$, where $-e$ is the electron's charge (so that $e > 0$), and P the orbital period. In terms of the orbital radius r and the orbital velocity v, we have $P = 2\pi r/v$, and the electron's angular momentum is $L = m_e v r$, where m_e is the electron mass. The area enclosed by the loop is πr^2, so that

$$\boldsymbol{\mu} = \left(-\frac{ev}{2\pi r}\right)(\pi r^2)\,\boldsymbol{n} = -\frac{evr}{2}\,\boldsymbol{n} = -\frac{e}{2m_e}\,L\boldsymbol{n} = -\frac{e}{2m_e}\,\boldsymbol{L}.$$

In this classical model, the molecular magnetic moment is proportional to the electron's orbital angular momentum.

In a multi-electron system, the contributions from each electron add vectorially, and $\boldsymbol{\mu} = -e\boldsymbol{L}/2m_e$, where $\boldsymbol{L}$ is now the *total* orbital angular momentum. This result would also be valid quantum mechanically, were it not for the electron's spin. Instead, the quantum mechanical result is

$$\boldsymbol{\mu} = -\frac{e}{2m_e}(\boldsymbol{L} + 2\boldsymbol{S}),$$

where $\boldsymbol{L}$ (total orbital angular momentum) and $\boldsymbol{S}$ (total spin angular momentum) are now quantum *operators*. These act as follows on their respective eigenstates:

$$
\begin{array}{llll}
\boldsymbol{L}^2|\ell, m_\ell\rangle &=& \ell(\ell+1)\hbar^2|\ell, m_\ell\rangle, & \ell = 0, 1, 2, \cdots \\
L_z|\ell, m_\ell\rangle &=& m_\ell\hbar|\ell, m_\ell\rangle, & m_\ell = -\ell, \cdots, \ell \\
\boldsymbol{S}^2|s, m_s\rangle &=& s(s+1)\hbar^2|s, m_s\rangle, & s = 0, \frac{1}{2}, 1, \cdots \\
S_z|s, m_s\rangle &=& m_s\hbar|s, m_s\rangle, & m_s = -s, \cdots, s
\end{array}
$$

The *total* angular momentum, $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, operates in a similar way:

$$
\begin{array}{llll}
\boldsymbol{J}^2|j, m_j\rangle &=& j(j+1)\hbar^2|j, m_j\rangle, & j = |l - s|, \cdots, l + s \\
J_z|j, m_j\rangle &=& m_j\hbar|j, m_j\rangle, & m_j = -j, \cdots, j
\end{array}
$$

It can be shown that $\boldsymbol{L} + 2\boldsymbol{S} = g\boldsymbol{J}$, where

$$g = \frac{3}{2} + \frac{s(s+1) - \ell(\ell+1)}{2j(j+1)}$$

is known as the *Landé factor*; it describes how a molecule's total angular momentum can be decomposed into orbital and spin parts. The molecular magnetic moment can therefore be expressed as

$$\boxed{\boldsymbol{\mu} = -g\mu_B\left(\frac{\boldsymbol{J}}{\hbar}\right)}. \tag{5.1}$$

The quantity

$$\boxed{\mu_B = \frac{e\hbar}{2m_e}} \tag{5.2}$$

is called the *Bohr magneton*; it represents the quantum unit of molecular magnetic moments. Different molecules have different total angular momenta, and different Landé factors; they therefore have different magnetic moments.

5.1.4   Forms of magnetism

A solid is immersed in an external magnetic field $\boldsymbol{B}_{\text{ext}}$. If the solid is magnetic, then the total field $\boldsymbol{B}_{\text{in}}$ *inside* the solid is different from $\boldsymbol{B}_{\text{ext}}$:

$$\boxed{\boldsymbol{B}_{\text{in}} = \boldsymbol{B}_{\text{ext}} + \mu_0 \boldsymbol{M}(T)}\,, \qquad (5.3)$$

where $\boldsymbol{M}(T)$ is the solid's *magnetization*, which depends on temperature. The magnetic response of a solid can also be described by its *susceptibility*,

$$\chi(T, B_{\text{ext}}) = \left(\frac{\partial M}{\partial B_{\text{ext}}} \right)_T,$$

where $M \equiv |\boldsymbol{M}|$ is the magnitude of the magnetization vector.

The different forms of magnetism have to do with how $\boldsymbol{M}$ and χ behave for different solids:

- **Diamagnetism:** In a diamagnetic solid, $\boldsymbol{M}$ points in the direction opposite to $\boldsymbol{B}_{\text{ext}}$, so that $\boldsymbol{B}_{\text{in}} < \boldsymbol{B}_{\text{ext}}$. Diamagnetic solids are made of molecules with vanishing total angular momentum, and therefore, vanishing magnetic moments. Diamagnetism is caused by atomic currents generated by magnetic induction; it is a very weak effect. Examples of diamagnetic materials are H_2O, Cu, Pb, and NaCl.

- **Paramagnetism:** In a paramagnetic solid, $\boldsymbol{M}$ points in the same direction as $\boldsymbol{B}_{\text{ext}}$, so that $\boldsymbol{B}_{\text{in}} > \boldsymbol{B}_{\text{ext}}$. Paramagnetic solids are made of molecules with nonzero total angular momentum, and therefore, nonzero magnetic moments. Paramagnetism is caused by the individual magnetic moments of all the molecules. For a unit volume containing N molecules, the magnetization is given by

$$\boxed{\boldsymbol{M}(T) = \left\langle \sum_{n=1}^{N} \boldsymbol{\mu}_n \right\rangle}\,, \qquad (5.4)$$

  where $\boldsymbol{\mu}_n$ is the magnetic moment of the nth molecule. (The average $\langle\ \rangle$ will be defined precisely below; it involves both quantum and thermal averaging.) In the absence of an applied field, the magnetic moments are randomly oriented, and $\boldsymbol{M} = 0$. In the presence of an applied field, however, the magnetic moments align themselves with $\boldsymbol{B}_{\text{ext}}$, and $\boldsymbol{M}$ therefore points in the direction of $\boldsymbol{B}_{\text{ext}}$. Only the interactions between the magnetic moments and the external field, through the Hamiltonian $H_n = -\boldsymbol{\mu}_n \cdot \boldsymbol{B}_{\text{ext}}$, are important in paramagnetic solids; mutual interactions among the magnetic moments are negligible. Paramagnetism is a stronger effect than diamagnetism. Examples of paramagnetic materials are Na, Al, $CuCl_2$, $NiSO_4$, and O_2.

- **Ferromagnetism:** Ferromagnetism is essentially like paramagnetism, except that mutual interactions among the magnetic moments are very important and cannot be neglected. These interactions can cause the magnetization to be permanent. Ferromagnetism is a much stronger effect than paramagnetism. Examples of ferromagnetic materials are Fe and Fe_3O_4.

5.2   General theory of paramagnetism

While ferromagnetism is the most interesting form of magnetism (that associated with permanent magnets), the importance of the mutual interactions among the

magnetic moments makes it quite difficult to model and analyze. On the other hand, because these interactions are negligible in the case of paramagnetism, this weaker form of magnetism lends itself much more easily to modeling and analysis. In this section, we formulate the general theory of paramagnetism.

5.2.1   The model

We imagine that a paramagnetic solid of unit volume contains N molecules which are located at specific sites on a lattice, so that they can be treated as distinguishable. We assume that the solid is immersed in a heat reservoir at a temperature T, so that it follows the Boltzmann distribution of probabilities. (We say that the system is in a canonical ensemble.)

We denote the total angular momentum of the n-th molecule by $\boldsymbol{J}_n$, and its magnetic moment is given by

$$\boldsymbol{\mu}_n = -g\mu_B(\boldsymbol{J}_n/\hbar),$$

where g is the Landé factor, and μ_B the Bohr magneton. We assume that g and j are the same for each molecule.

The molecules are assumed to interact only with an external magnetic field $\boldsymbol{B}$, and not with each other. The kinetic energy of the molecules is also neglected. The Hamiltonian describing the system is therefore

$$\boxed{H = -\sum_{n=1}^{N} \boldsymbol{\mu}_n \cdot \boldsymbol{B}} \; . \tag{5.5}$$

Finally, the magnetization is in the direction of $\boldsymbol{B}$, which we choose to be the z-direction. The magnetization is therefore given by $\boldsymbol{M} = M\hat{\boldsymbol{z}}$, where $M = \langle \sum_{n=1}^{N} \mu_{nz} \rangle$, or

$$\boxed{M = -g\mu_B \left\langle \sum_{n=1}^{N} \frac{J_{nz}}{\hbar} \right\rangle} \; . \tag{5.6}$$

The magnetization will be the thermodynamic quantity of prime interest in the following. To calculate it we need to spell out how $\sum_{n=1}^{N} J_{nz}$ is to be averaged.

5.2.2   Quantum and thermal averages

The magnetization of a paramagnetic solid is a macroscopic quantity, and the purpose of the averaging $\langle \; \rangle$ is to go from the microscopic quantities J_{nz} to something which does not depend on the detailed properties of the solid. (This rejection of the microscopic details to define macroscopic quantities is of course the central theme of statistical mechanics.) The magnetization, therefore, is a classical quantity, which means that unlike J_{nz}, it cannot be a quantum operator. To produce such a classical quantity, the averaging $\langle \; \rangle$ must therefore incorporate a *quantum averaging*, which we now describe.

Let $|r\rangle$ denote one of the possible quantum states of the entire system, which consists of N distinguishable molecules. To characterize this quantum state, we note that the nth molecule must be in a simultaneous eigenstate of $\boldsymbol{J}_n^2$ and J_{nz}, which we denote $|j, m_n\rangle$ or simply $|m_n\rangle$, because j is the same for all the molecules. To know the individual state of every molecule is to know the state of the entire system, and therefore, this state is fully determined once all N values of the quantum numbers m_n are specified. Mathematically,

$$r = \{m_1, m_2, \cdots, m_N\} \equiv \{m_n\},$$

and we may express $|r\rangle$ as the tensor product of all the individual states $|m_n\rangle$:

$$|r\rangle = |m_1\rangle|m_2\rangle \cdots |m_N\rangle.$$

The *quantum average* of some operator $\mathcal{O}$ consists in taking its expectation value in the quantum state $|r\rangle$:

$$\langle r|\mathcal{O}|r\rangle \equiv \text{quantum average of operator } \mathcal{O}.$$

Unlike $\mathcal{O}$, $\langle r|\mathcal{O}|r\rangle$ is not a quantum operator, and it is therefore suitable to represent a classical object.

Part of the averaging $\langle\ \rangle$ therefore consists in performing a quantum average. Now, if we knew the exact quantum state of the solid, then this would be the only form of averaging that would need to be considered. But this is only part of the story, because in reality, the quantum state is not only unknown, it is subject to fluctuations due to the solid's interaction with the heat reservoir. The averaging $\langle\ \rangle$ must therefore also incorporate a *thermal averaging* over the Boltzmann distribution of probabilities.

The (quantum-thermal) average $\langle\mathcal{O}\rangle$ of the operator $\mathcal{O}$ is therefore defined by

$$\boxed{\langle\mathcal{O}\rangle = \overline{\langle r|\mathcal{O}|r\rangle} = \sum_r \langle r|\mathcal{O}|r\rangle\, p_r = \frac{1}{Z} \sum_r \langle r|\mathcal{O}|r\rangle\, e^{-\beta E_r}} , \qquad (5.7)$$

where Z is the partition function, $\beta = 1/kT$, and E_r is the energy eigenvalue of the quantum state $|r\rangle$.

To calculate M we must therefore determine the energy eigenvalues, evaluate the partition function, and then carry out the sum over states that appears at the right of Eq. (5.7). We will consider each of these steps in turn.

5.2.3   Energy eigenvalues

The energy eigenvalue of the quantum state $|r\rangle$ is obtained by acting on $|r\rangle$ with the Hamiltonian operator of Eq. (5.5). Since $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, we have

$$H = -B \sum_{n=1}^{N} \mu_{nz} = g\mu_B B(J_{1z}/\hbar + J_{2z}/\hbar + \cdots + J_{Nz}/\hbar).$$

How does this operate on $|r\rangle = |m_1\rangle|m_2\rangle \cdots |m_N\rangle$? The answer is that when J_{1z} acts on $|r\rangle$, it notices only the presence of $|m_1\rangle$ and leaves the rest of $|r\rangle$ untouched. Since $J_{1z}|m_1\rangle = m_1\hbar|m_1\rangle$, we find that $J_{1z}|r\rangle = m_1\hbar|r\rangle$. Proceeding similarly with the other molecules, we obtain

$$H|r\rangle = g\mu_B B(m_1 + m_2 + \cdots m_N)|r\rangle.$$

This is clearly of the form $H|r\rangle = E_r|r\rangle$, and we conclude that the energy eigenvalues are given by

$$\boxed{E_r = g\mu_B B \sum_{n=1}^{N} m_n} . \qquad (5.8)$$

5.2.4   Partition function

The partition function is obtained by summing the Boltzmann factor $e^{-\beta E_r}$ over all the quantum states $|r\rangle$. More concretely, this means that we must sum $e^{-\beta E_r}$ over

all the values that each individual quantum number m_n can take. In other words,

$$
\begin{aligned}
Z &= \sum_{m_1}\sum_{m_2}\cdots\sum_{m_N}\exp\left[-\beta g\mu_B B(m_1 + m_2 + \cdots + m_N)\right] \\
&= \left[\sum_{m_1}\exp(-\beta g\mu_B B m_1)\right]\left[\sum_{m_2}\exp(-\beta g\mu_B B m_2)\right]\cdots\left[\sum_{m_N}\exp(-\beta g\mu_B B m_N)\right]
\end{aligned}
$$

Since the sums are all identical, this simplifies to

$$
Z = \left[\sum_{m=-j}^{j}\exp\left(-\beta g\mu_B B m\right)\right]^N, \tag{5.9}
$$

where we have indicated that each m_n varies from $-j$ to $+j$ in integral steps. We will evaluate this sum explicitly below.

5.2.5   Magnetization

We may now calculate the magnetization. We start by combining Eqs. (5.6) and (5.7), and we proceed by going through some of the same steps as for the calculation of the energy eigenvalues. After a short computation, we find that the magnetization is given by

$$
\begin{aligned}
M &= -\frac{g\mu_B}{Z}\sum_{r}\left(\sum_{n=1}^{N}m_n\right)e^{-\beta E_r} \\
&= -\frac{g\mu_B}{Z}\sum_{m_1}\cdots\sum_{m_N}(m_1 + \cdots + m_N)\exp\left[-\beta g\mu_B B(m_1 + \cdots + m_N)\right].
\end{aligned}
$$

We now employ the familiar trick of expressing a thermodynamic quantity of interest — in this case the magnetization — directly in terms of the partition function. It should be clear that our expression for M differs from

$$
Z = \sum_{m_1}\cdots\sum_{m_N}\exp\left[-\beta g\mu_B B(m_1 + \cdots + m_N)\right]
$$

only by the presence of the factor $-g\mu_B/Z$ in front of the sums, and the factor $(m_1 + \cdots + m_N)$ inside the sums. This last factor can be generated by differentiating Z with respect to B, say. Since this also produces the factor $-\beta g\mu_B$, we are not far from the desired result, which is

$$
M = \frac{1}{\beta}\frac{\partial \ln Z}{\partial B}. \tag{5.10}
$$

The magnetization of a paramagnetic solid is therefore calculated in two steps. First, the partition function is computed from Eq. (5.9). Second, M is obtained from Eq. (5.10). By themselves, these two equations encapsulate the theory of paramagnetism.

5.3   Molecules with $j = \frac{1}{2}$

To illustrate how the general theory is applied, we consider in this section the simple (but unrealistic) case of molecules with spin 1/2. We therefore set the orbital angular momentum to zero, so that $\boldsymbol{J} = \boldsymbol{S}$, with $s = 1/2$. This gives us a Landé factor $g = 2$.

Because m can take the values $\pm\frac{1}{2}$ only, Eq. (5.9) gives $Z = [\exp(\beta\mu_B B) + \exp(-\beta\mu_B B)]^N$, or

$$Z = \Big[2\cosh(\beta\mu_B B)\Big]^N.$$

This implies that $\ln Z = N\ln[2\cosh(\beta\mu_B B)]$, and the magnetization follows by partial differentiation with respect to B. We easily obtain

$$M = N\mu_B \tanh(\beta\mu_B B).$$

We see that the magnetization is equal to $N\mu_B$ times a function of $\beta\mu_B B = \mu_B B/kT$, the ratio of the magnetic unit of energy to the thermal unit of energy.

It is interesting to ask how the magnetization behaves at low temperatures, when the thermal unit of energy is much smaller than the magnetic unit of energy. In this limit, the argument of the hyperbolic tangent is very large, and since $\tanh(x \ll 1) \sim 1$, we have that

$$M \sim N\mu_B, \qquad kT \ll \mu_B B.$$

This result means that in the absence of thermal agitations, the magnetic moments are all aligned with the external magnetic field, thus minimizing the magnetic potential energy. Since the individual moments are all of magnitude μ_B, by virtue of Eq. (5.1), and since they all point in the same direction, the magnetization must clearly be equal to $N\mu_B$, which is indeed what we find.

In the limit of high temperatures, when the thermal unit of energy is much larger than the magnetic unit of energy, the situation is completely different. Here, the thermal agitations are so important that the magnetic moments are prevented from aligning themselves with the external magnetic field. Instead, the moments point in more or less random directions, and as a result, the magnetization is very small. Quantitatively, we may use the result $\tanh(x \ll 1) \sim x$ to derive

$$M \sim N\mu_B \frac{\mu_B B}{kT}, \qquad kT \gg \mu_B B.$$

The statement that $M \propto 1/T$ at large temperatures is known as *Curie's law*.

5.4   Molecules with arbitrary j

It is not too difficult to calculate the magnetization when j is arbitrary. Although the mathematics are slightly harder than in the simplest case $j = 1/2$, the basic physics is exactly the same.

We must first evaluate the sum $Z_1 = \sum_{m=-j}^{j} e^{-2\eta m}$, which gives the partition function for a single molecule. For convenience, we have defined the quantity

$$\boxed{\eta = \frac{g\mu_B B}{2kT}}, \tag{5.11}$$

the ratio of magnetic and thermal units of energy. How do we evaluate this sum over m? To answer this, consider instead the quantity $(1 - e^{2\eta})Z_1$, which is given by

$$
\begin{aligned}
(1 - e^{2\eta})Z_1 \quad = \quad & e^{2\eta j} + e^{2\eta(j-1)} + \cdots + e^{-2\eta(j-1)} + e^{-2\eta j} \\
& - e^{2\eta(j+1)} - e^{2\eta j} - e^{2\eta(j-1)} - \cdots - e^{-2\eta(j-1)}.
\end{aligned}
$$

We see that most of the terms cancel out, leaving

$$Z_1 = \frac{e^{-2\eta j} - e^{2\eta(j+1)}}{1 - e^{2\eta}}.$$

Multiplying above and below by $e^{-\eta}$ yields

$$Z_1 = \frac{\sinh\big[(2j+1)\eta\big]}{\sinh(\eta)}.$$

Finally, raising this to the Nth power gives the full partition function:

$$Z = \left\{ \frac{\sinh\big[(2j+1)\eta\big]}{\sinh(\eta)} \right\}^N. \tag{5.12}$$

It is a good exercise to check that this reduces to the previous result when $j = 1/2$.

The magnetization is now obtained by partial differentiation, using Eq. (5.10). The algebraic steps are easy, and we obtain

$$M = jgN\mu_B\, B_j(\eta), \tag{5.13}$$

where $B_j(\eta)$, the *Brillouin function*, is given by

$$B_j(\eta) = \frac{1}{2j}\left\{ (2j+1)\coth\big[(2j+1)\eta\big] - \coth(\eta) \right\}. \tag{5.14}$$

A plot of the Brillouin function, for $j = \{\frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}\}$, appears on the side of this page. It may easily be shown that $B_j(\eta \ll 1) \sim \frac{2}{3}(j+1)\eta$, while $B_j(\eta \gg 1) \sim 1$.

These limiting expressions for the Brillouin function can be used to derive the following results. First, at low temperatures $(\eta \gg 1)$, the magnetization is given by

$$M \sim jgN\mu_B, \qquad kT \ll \mu_B B. \tag{5.15}$$

The physical picture is the same as before: at low temperatures, the magnetic moments are all aligned with the external magnetic field, the potential energy is minimized, and the magnetization is maximized. Second, at high temperatures $(\eta \ll 1)$, the magnetization is now given by

$$M \sim \tfrac{1}{3}j(j+1)g^2\, N\mu_B\, \frac{\mu_B B}{kT}, \qquad kT \gg \mu_B B. \tag{5.16}$$

Again the interpretation is as before: at high temperatures the magnetic moments are randomly oriented, and the magnetization is very small. Equation (5.16) is the most general form of Curie's law.

Equations (5.13) and (5.14) are our theoretical predictions for the magnetization of a paramagnetic solid. The natural question, of course, is how well do these predictions compare with experimental data? The answer is very well indeed, as a quick glance at p. 262 of Reif's book will confirm. This agreement is quite remarkable, given the extreme simplicity of our model.

5.5 Paramagnetic solid in isolation

In Sec. 2 we have formulated the general theory of paramagnetism under the assumption that the paramagnetic solid was immersed in a heat bath at a temperature T. In this section we reconsider this problem, this time under the assumption that the paramagnetic solid is isolated. The reason for this is twofold. First, this will provide a nice illustration of the fact that the predictions of statistical mechanics do not typically depend on whether the system is assumed to be isolated or in interaction with a heat bath. Second, and more important, we will be led toward the striking conclusion that the temperature of an isolated system can sometimes be negative! Throughout this section, to simplify the discussion as much as possible, it will be assumed that the paramagnetic solid is composed of molecules with $j = \frac{1}{2}$.
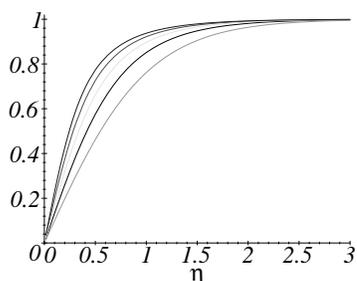
5.5.1   Statistical weight and entropy

We consider an isolated system consisting of N molecules, each of which is located on a specific site on a lattice. The nth molecule has a magnetic moment $\mu_n = -2\mu_B m_n$, where m_n is capable of taking the values $\pm\frac{1}{2}$. The energy of this molecule is given by $\varepsilon_n = 2\mu_B B m_n$. We want to study the thermodynamic properties of this isolated system. From what we've learned in part B of this course, we know that we must first compute $\Omega(E)$, the statistical weight, for this system.

If we denote by n the number of molecules with $m_n = +\frac{1}{2}$, so that their magnetic moments are negative and their energies positive, then the magnetization of the solid will be given by $M = n(-\mu_B) + (N - n)(+\mu_B)$, or

$$M = (N - 2n)\mu_B.$$

Similarly, the total energy of the solid will be

$$E = (2n - N)\mu_B B = -MB.$$

Of course, we do not know *which*, among the N molecules, are those with a negative magnetic moment. We therefore want to calculate the number of ways the system can be prepared so that a number n of molecules will have a negative magnetic moment. Because E varies with n and nothing else (because N, μ_B, and B are all constants), this is the same as $\Omega(E)$, the number of ways the system can be prepared so that it will have an energy E. The hard work of calculating this quantity was done in a homework problem, where it was found that

$$\Omega(E) = \frac{N!}{n!(N - n)!}.$$

Here, n is given in terms of E by

$$n = \frac{1}{2}(N + E/\varepsilon),$$

where we have defined the quantity $\varepsilon \equiv \mu_B B$ for convenience. Notice that $n < N/2$ if the energy is negative, while $n > N/2$ if the energy is positive. Notice also that $\Omega = 1$ if either $n = 0$ or $n = N$. This means that the entropy is zero whenever the energy reaches the limiting values $\pm N\varepsilon$.

An explicit expression for the entropy follows from the relation $S = k \ln \Omega$. Assuming that N, n, and $N - n$ are all large, and using the Stirling approximation $\ln(M!) \simeq M \ln M - M$ for large M, we find

$$S(E) = k\big[N \ln N - n \ln n - (N - n) \ln(N - n)\big].$$

This expression confirms that $S = 0$ when $E = \pm N\varepsilon$.

5.5.2   Temperature

The temperature of an isolated system is defined by recalling the first law of thermodynamics, $dE = T\,dS + \cdots$, from which it follows that

$$\beta \equiv \frac{1}{kT} = \frac{1}{k}\left(\frac{\partial S}{\partial E}\right)_{V,N}.$$

Taking into account that S depends on E through the variable n, the partial derivative is easily evaluated, and we find

$$\beta = \frac{1}{2\varepsilon} \ln\left(\frac{N\varepsilon - E}{N\varepsilon + E}\right).$$
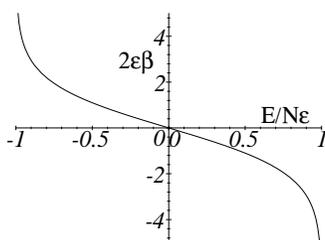
The function $\beta(E)$ is plotted on the side of the page. Let us now analyze the behaviour of this function.

In the limit $E \to -N\varepsilon$, which corresponds to the magnetic moments all pointing in the direction of the external magnetic field, $\beta \to +\infty$, which corresponds to $T \to 0^+$. This makes good sense: the temperature must be very small if the magnetic moments can all align themselves with the external field.

As the energy increases from $-N\varepsilon$, β decreases, and the temperature increases. When $E \to 0^-$, which corresponds to an equal number of aligned and anti-aligned magnetic moments, $\beta \to 0^+$, which means that $T \to +\infty$. This also makes sense: the temperature must be very large for the magnetic moments to be randomly oriented.

Now for the big surprise. The system's energy is not required to be smaller than zero. If a larger number of magnetic moments are pointing in the direction opposite to the magnetic field, then $n > N/2$, and $E > 0$. Now, β is perfectly well behaved in the neighbourhood of $E = 0$. As E goes from 0^- to 0^+, β simply goes from 0^+ to 0^-. As E is increased further, β grows more negative. Finally, as E reaches the upper limit $+N\varepsilon$, which corresponds to the magnetic moments all anti-aligned with the external field, $\beta \to -\infty$.

Of course, this innocuous behaviour of β translates into a rather strange behaviour for the temperature itself. As β gently moves from 0^+ to 0^-, T jumps wildly from $+\infty$ to $-\infty$! And as β grows more negative as the energy is further increased, T increases from $-\infty$ until it reaches 0^-! Strange behaviour indeed. Strange also is the following conclusion: Because a body containing more energy must be hotter than the same body with less energy, a body at a negative temperature must be hotter than a body at a positive temperature!

5.5.3   Negative temperatures?

Do negative temperatures really make physical sense? To answer this question, let us first ask why the temperature can go negative in a system of noninteracting magnetic moments. The answer, as we shall explain in the following two paragraphs, is this: The key point with this system is that its energy has an *upper bound*, given here by $+N\varepsilon$. This is a very uncommon situation, because the existence of such an upper bound is clearly impossible if the system has kinetic energy. In fact, the prediction of negative temperatures in isolated paramagnetic solids has a lot to do with the fact that we have ignored, all along, the kinetic energy of the molecules within the solid. This is not a bad approximation at sufficiently small (positive) temperatures, but the kinetic energy would have to be taken into account at large temperatures. Doing so would then reveal that negative temperatures do not occur.

How does the existence of an upper bound in the system's energy imply the occurrence of negative temperatures? To answer this, consider a system who's energy is bounded below by $E_{\min}$, and above by $E_{\max}$. When $E = E_{\min}$, the system is in a very special state, the ground state. (For example, in a paramagnetic solid, every molecule must have its magnetic moment aligned with the external magnetic field.) Because the ground state is unique, the statistical weight is equal to unity when $E = E_{\min}$, and it increases as E increases away from $E_{\min}$. On the other hand, the system must also be in a very special state when $E = E_{\max}$. (For example, in a paramagnetic solid, every molecule must have its magnetic moment anti-aligned with the external magnetic field.) Because this state is just as unique as the ground state, the statistical weight must also be equal to unity when $E = E_{\max}$. After having increased for a while, the statistical weight must eventually decrease back to unity as E approaches $E_{\max}$. The entropy must therefore have the following behaviour: it increases from zero when E is increased from $E_{\min}$, and it eventually decreases to zero when E approaches $E_{\max}$; somewhere in between, $S(E)$ must

reach a maximum.

   The occurrence of negative temperatures follows immediately from the last statement. For isolated systems, $1/T$ is *defined* to be $\partial S/\partial E$, the slope of the entropy function. This slope is clearly positive when E is near $E_{\min}$, but it is clearly negative when E approaches $E_{\max}$. At the point where $S(E)$ reaches its maximum, the slope goes to zero, and $1/T$ changes sign.

   Although negative temperatures do not naturally occur, they can nonetheless be simulated in the laboratory. In 1951, Purcell and Pound prepared a system of nuclear spins in such a way that initially, most of the spins were aligned with an external magnetic field. By suddenly reversing the direction of the field, they were able to momentarily create a system of anti-aligned spins, to which a negative temperature must be assigned. This system is short lived, because the spins eventually re-align themselves with the external field. However, the timescale for this to occur is sufficiently long that for good while, the system can be considered to be in equilibrium at a negative temperature.

5.5.4   Correspondence with the canonical ensemble

We conclude this section by showing how the results obtained here (for positive temperatures!) agree with those of Sec. 3. Our previous relation for β can easily be inverted to give

$$E = -N\varepsilon \tanh(\beta\varepsilon).$$

Using the fact that $M = -E/B$ and replacing ε by $\mu_B B$, we arrive at

$$M = N\mu_B \tanh(\beta\mu_B B),$$

which is exactly as in Sec. 3. This provides a nice illustration of the general idea that the physical predictions of statistical mechanics do not typically depend on the choice of probability distribution. In this case, the predictions based upon the uniform distribution (appropriate for an isolated system) turn out to be identical to those based upon the Boltzmann distribution (appropriate for a system interacting with a heat reservoir).

5.6   Problems

1. Using the general theory, show that the mean energy E and the magnetization M of a paramagnetic solid are related by

   $$E = -MB,$$

   where B is the applied magnetic field.

2. (Rief 7.14) Consider a unit volume containing N magnetic atoms at a temperature T, and describe the situation *classically*. Then each magnetic moment $\vec{\mu}$ can make any angle θ with respect to a given direction (call it the z direction).

   In the absence of an applied magnetic field, the probability that this angle lies between θ and $\theta + d\theta$ is proportional to the solid angle $2\pi \sin\theta \, d\theta$ enclosed in this range. In the presence of an applied magnetic field B in the z direction, this probability must further be proportional to the Boltzmann factor $e^{-U/kT}$, where U is the potential energy of the moment $\vec{\mu}$ making this angle θ with the z direction.

   a) Use this result to calculate the classical expression for the mean value of μ_z, the magnetic moment of a single atom along the z direction. This is the mean magnetic moment of a *single* atom. Multiplying this by N then gives the classical expression for the magnetization of the material.

b) This classical analysis should be valid for high temperatures, that is, for temperatures T such $kT \gg \mu B$. Show that in this limit, your expression for part a) reduces to

$$M = \frac{N\mu^2 B}{3kT}.$$

This is yet another form of Curie's law.

3. Atoms with total angular momentum $j = \frac{1}{2}$ and Landé factor $g = 2$ at temperature T align themselves either parallel or antiparallel to the direction of an applied magnetic field. Determine the respective fractions of atoms aligned parallel and antiparallel to a field of 10.0 Tesla at a temperature of 10.0 K. (You must treat this problem quantum mechanically.)

4. (Mandl 3.4) The magnetic moment μ of an atomic nucleus is of order

$$\frac{Ze\hbar}{2Am_p},$$

where m_p is the proton mass, Z the atomic number, and A the mass number. Most nuclei have $A \simeq 2Z$. Estimate the magnetic field required at a temperature of 0.01 K to produce an appreciable alignment of the nuclei along the direction of the field.

CHAPTER 6
QUANTUM STATISTICS OF IDEAL GASES

In this section we return to a familiar system: the ideal gas. But while our previous considerations were limited to the classical limit, in this section we will study the ideal gas at a much more fundamental level that includes in full the effects of quantum mechanics. New phenomena will thus be brought to light.

Before getting on with our discussion of the quantum statistics of ideal gases, however, it will be necessary to sharpen our statistical-mechanics tools so as to make them easier to use. We will thus be led toward a new formulation of statistical mechanics, one which is entirely equivalent to the one encountered before. A major advantage of this new formulation will be the removal of any danger of over-counting the number of states by not taking into account the indistinguishability of the particles. You may recall that this problem was overcome in Sec. B 6 by the *ad hoc* procedure of dividing the statistical weight by $N!$, the total number of permutations among the N particles. In our new formulation, the indistinguishability of identical particles in quantum mechanics will be taken into account from the very start.

This reformulation will take a while to get accomplished, and during this time, our considerations will be largely formal. We will derive results that will appear to be very abstract, and the advantages of doing all this will not be clear immediately. Please bear with it! The payoff will come.

6.1 Quantum statistics

We want to formulate, in a very precise manner, the statistical mechanics of a system of N identical, noninteracting particles confined to a volume V and in thermal contact with a heat bath at a temperature T. The relevant statistical ensemble for this problem is the canonical ensemble, whose most important quantity is the partition function $Z(\beta, V, N)$, where $\beta = 1/kT$. To compute this we must sum the Boltzmann factor $e^{-\beta E}$ over all states accessible to the entire system. The major difficulty in doing this is to specify what these states are, in a way that takes into account the fact that the particles are indistinguishable. This is where our reformulation of statistical mechanics comes in.

6.1.1   Microstates

Instead of thinking about the states of the *entire system*, focus for the time being on the states accessible to a *single particle*. These will be called *single-particle states*, and they are extremely easy to characterize: We already know that these states are labeled by three integer-valued quantum numbers, n_x, n_y, and n_z. This, however,

75

is only a partial description, because it ignores the *internal* states that arise because of particle spin. For example, an electron has a spin $s = \frac{1}{2}$, which can have either one of the two projections $m_s = \pm\frac{1}{2}$. The full characterization of the single-particle states therefore requires the specification of four quantum numbers, n_x, n_y, n_z, and m_s.

We may introduce the abstract index r to label the single-particle states. The index r stands for the set of all four quantum numbers that are necessary to completely specify the states:

$$r = \{n_x, n_y, n_z, m_s\}.$$

There is obviously an infinite number of single-particle states, because n_x, n_y, and n_z can all take an infinite number of values. The quantum number m_s, on the other hand, ranges from $-s$ to $+s$ with integral steps. The zero-energy state $(n_x, n_y, n_z, m_s) = (0, 0, 0, -s)$ will be symbolically represented by $r = 0$. Similarly, $r = 1$ stands for $(0, 0, 0, -s + 1)$, and so on.

The energy eigenvalue corresponding to the single-particle state r is denoted ε_r. We know that this is given by

$$\varepsilon_r = \frac{(2\pi\hbar)^2}{8mV^{2/3}} \left(n_x{}^2 + n_y{}^2 + n_z{}^2\right),$$

which shows that ε_r is independent of m_s.

How do we go from the single-particle states to the states of the entire system? The key is to introduce the quantity n_r, called the *occupation number* of the single-particle state r. By definition, this is the total number of particles which happen to occupy the particular single-particle state r. Notice that we do not care *which* particles among the N particles happen to occupy the specific state r. All we are asking is *how many* particles do, and the answer is provided by the occupation number n_r. (This is where the indistinguishability of the particles is taken into account.) Once we have the complete list of all the occupation numbers, we know precisely how many particles occupy each and everyone of the single-particle states r. Since this is all that can be known about the system without violating the indistinguishability postulate of quantum mechanics, the state of the entire system is now completely determined. If we use the abstract index R to label the states of the entire system, then R stands for the complete set of all the occupation numbers, one for each of the single-particle states r. In mathematical notation,

$$R = \{n_0, n_1, n_2, \ldots\} = \{n_r\}.$$

We now have a way of characterizing the states of the entire system in terms of the single-particle states.

The total energy of the system is given by $\epsilon_0(= 0)$ times n_0, the number of particles in the state $r = 0$, plus ϵ_1 times n_1, the number of particles in the state $r = 1$, plus all other contributions. If we denote the total energy of the state R by E_R, then

$$E_R = \sum_r n_r\, \varepsilon_r,$$

where the sum is over all single-particle states. Similarly, the total number of particles in the state R is given by the sum of all the occupation numbers:

$$N_R = \sum_r n_r.$$

In our case, the number of particles is fixed, so $N_R = N$ for all states R. This shows that the occupation numbers are not all independent, because their sum must always reproduce the total number of particles.

6.1.2   Bosons and fermions

It is one of the most fundamental facts of nature that elementary particles, atoms, and molecules are divided into two groups called bosons and fermions. Bosons are particles possessing integer spin ($s = 0, 1, 2, \cdots$). For the purpose of this section, the most important property of bosons is that an *arbitrary* number of them can exist in the same single-particle state r. In other words, their occupation numbers can take any value between 0 and N:

$$n_r = 0, 1, 2, \cdots, N \qquad \text{(bosons)}.$$

Fermions, on the other hand, are particles possessing half-integer spin ($s = \frac{1}{2}, \frac{3}{2}, \cdots$). The most important property of fermions is that two or more fermions *cannot* occupy the same single-particle state r. Their occupation numbers, therefore, can only take the values 0 and 1:

$$n_r = 0, 1 \qquad \text{(fermions)}.$$

For both bosons and fermions, the occupation numbers are further restricted by the condition $\sum_r n_r = N$.

6.1.3   Partition function

Let us now attempt to calculate the partition function of an ideal gas consisting of bosons. Using our previous results, the partition function is given by

$$Z = \sum_R e^{-\beta E_R} = \sum_{\{n_r\}} \exp\left(-\beta \sum_r n_r \varepsilon_r\right),$$

where the sum is over all the individual occupation numbers n_r. Because of the restriction on $\sum_r n_r$, this becomes

$$Z = \sum_{n_0=0}^{N} e^{-\beta n_0 \varepsilon_0} \sum_{n_1=0}^{N-n_0} e^{-\beta n_1 \varepsilon_1} \sum_{n_2=0}^{N-n_0-n_1} e^{-\beta n_2 \varepsilon_2} \times \cdots.$$

These sums are very difficult to evaluate, because of the explicit dependence of the upper limits on the occupation numbers. There seems to be little hope that this calculation could be pushed forward.

Evaluation of the partition function is made difficult by the constraint $\sum_r n_r = N$, which enforces the condition that we are dealing with a fixed number of particles. If, however, this condition were to be removed, then the constraint would disappear, and the sums could presumably be evaluated much more easily. Could we possibly do this and get away with it?

6.1.4   Grand partition function

The answer is yes. Let us *pretend* that the number of particles is not fixed, and that the system's interaction with the reservoir allows for an exchange of particles. In effect, we are switching from a canonical (Boltzmann) description to a grand canonical (Gibbs) description. This switch should not produce any changes in our physical predictions: As we have seen in Sec. C 3, the fluctuation $\Delta N/N$ in the number of particles in a system described by the Gibbs distribution is extremely small, of the order of $N^{-1/2}$. So, although we shall no longer *formally* impose that the number of particles be fixed, this number will still be fixed for all practical purposes. The advantages associated with the removal of the constraint $\sum_r n_r = N$ will thus be obtained at virtually no cost at all.

The physical picture is now the following. We have a system of identical, non-interacting particles confined to a volume V and in contact with a reservoir at a temperature T and with a chemical potential μ. As a result of the interaction between the system and the reservoir, energy and particles are exchanged, but only in microscopic amounts, as the fluctuations are extremely small. The most important quantity of this description is $\mathcal{Z}(\alpha, \beta, V)$, the grand partition function, where $\alpha = -\mu/kT$, $\beta = 1/kT$. This is given by

$$\mathcal{Z} = \sum_R e^{-(\alpha N_R + \beta E_R)} = \sum_{\{n_r\}} \exp\left[-\left(\alpha \sum_r n_r + \beta \sum_r n_r \varepsilon_r\right)\right],$$

and becomes, after a few lines of algebra,

$$\mathcal{Z} = \sum_{n_0} e^{-(\alpha + \beta \varepsilon_0) n_0} \sum_{n_1} e^{-(\alpha + \beta \varepsilon_1) n_1} \sum_{n_2} e^{-(\alpha + \beta \varepsilon_2) n_2} \times \cdots . \tag{6.1}$$

We shall leave the grand partition function in this form for the time being. The main advantage of this expression over what we had previously obtained for Z is that the occupation numbers are now all independent. For this reason, the sums will be easy to evaluate.

This alternative description of our problem has introduced a new quantity, α, which is related to the chemical potential μ by $\alpha = -\mu/kT$. At the same time, the quantity N seems to have disappeared from the description. However, because of the fact that the fluctuations are small, we can identify N with $\overline{N_R}$, the system's mean number of particles. This is related to α through the relation (3.21), $\overline{N_R} = -\partial \ln \mathcal{Z}/\partial \alpha$. The parameter α is therefore determined implicitly in terms of N by the relation

$$N = -\frac{\partial \ln \mathcal{Z}}{\partial \alpha} . \tag{6.2}$$

6.1.5   Mean occupation numbers

We have seen in Sec. C 3 that once the grand partition function has been obtained, the thermodynamic quantities of the system follow immediately by straightforward manipulations. In the following it will be convenient to deal not with $\mathcal{Z}$ directly, but with the quantities $\bar{n}_r$, the *mean occupation numbers* of the single-particle states r. The reason is that these quantities are much simpler than the grand partition function itself. From the mean occupation numbers we will be able to directly compute the thermodynamic variables, without the need to compute $\mathcal{Z}$ explicitly.

The mean occupation number of a particular single-particle state r^* is obtained by averaging n_{r^*} over the Gibbs distribution:

$$\bar{n}_{r^*} = \frac{1}{\mathcal{Z}} \sum_R n_{r^*} e^{-(\alpha N_R + \beta E_R)},$$

where the sum is over all states accessible to the entire system. Writing this out more explicitly, we have

$$\bar{n}_{r^*} = \frac{1}{\mathcal{Z}} \sum_{\{n_r\}} n_{r^*} \exp\left\{-\left[\alpha(n_0 + \cdots + n_{r^*} + \cdots) + \beta(n_0 \varepsilon_0 + \cdots + n_{r^*} \varepsilon_{r^*} + \cdots)\right]\right\}.$$

But we see that the sum on the right-hand side is equal to $-\beta^{-1} \partial \mathcal{Z}/\partial \varepsilon_{r^*}$. We therefore have

$$\bar{n}_r = -\frac{1}{\beta} \frac{\partial \ln \mathcal{Z}}{\partial \varepsilon_r} \tag{6.3}$$

for any single-particle state r. This equation shows that the mean occupation numbers can all be obtained from the grand partition function. We will show the converse below.

6.1.6   Thermodynamic quantities

We now show how such thermodynamic quantities as the mean number of particles N, the mean energy E, and the pressure P can all be obtained directly in terms of the mean occupation numbers, without the need to first compute $\mathcal{Z}$.

Summing over all the mean occupation numbers should obviously return N, the mean number of particles. This is easily verified if we rewrite Eq. (6.2) in the form

$$N = -\frac{1}{\mathcal{Z}} \frac{\partial}{\partial \alpha} \sum_{\{n_r\}} \exp\left\{-\left[\alpha(n_0 + n_1 + \cdots) + \beta(n_0 \varepsilon_0 + n_1 \varepsilon_1 + \cdots)\right]\right\}.$$

Taking the derivative over α brings down the factor $-(n_0 + n_1 + \cdots)$, and we obtain $N = \bar{n}_0 + \bar{n}_1 + \cdots$. In other words,

$$\boxed{N = \sum_r \bar{n}_r} \, , \tag{6.4}$$

which confirms our initial expectation.

A similar calculation reveals that the mean energy E is obtained by summing $\bar{n}_r \, \varepsilon_r$ over all possible single-particle states:

$$\boxed{E = \sum_r \bar{n}_r \, \varepsilon_r} \, . \tag{6.5}$$

This result also appears quite reasonable.

The pressure follows easily. We recall from Eq. (3.24) that $P = kT \partial \ln \mathcal{Z}/\partial V$. Since the grand partition function depends on V only through the energy eigenvalues, $\varepsilon_r \propto V^{-2/3}$, we have that

$$P = kT \sum_r \frac{\partial \ln \mathcal{Z}}{\partial \varepsilon_r} \frac{\partial \varepsilon_r}{\partial V},$$

where we have applied the chain rule of differential calculus. Now, the first factor within the sum is equal to $-\beta \bar{n}_r$. The second factor is equal to $-2\varepsilon_r/3V$, and the previous equation becomes $P = (2kT\beta/3V) \sum_r \bar{n}_r \varepsilon_r$. Using Eq. (6.5), we arrive at our final result,

$$\boxed{P = \frac{2E}{3V}} \, . \tag{6.6}$$

Equations (6.1)–(6.6) are quite general: they are valid for *any* system of noninteracting particles, whether they are bosons or fermions. These equations, however, are still rather formal, and we cannot yet use them in practical calculations. To get there is the purpose of the following two sections.

6.2   Bose-Einstein statistics

6.2.1   Grand partition function

We now specialize to an ideal gas consisting of boson particles, and we want to calculate the grand partition function for this gas. We therefore return to Eq. (6.1),

and observe that $\mathcal{Z}$ is given by an infinite product of factors of the form

$$\sum_{n_r} e^{-(\alpha+\beta\varepsilon_r)n_r} = \sum_{n_r=0}^{\infty} \left[e^{-(\alpha+\beta\varepsilon_r)} \right]^{n_r},$$

where we have indicated that for bosons, there is no restriction on the value of the occupation numbers. This sum is easily evaluated, if we recall that for $x < 1$, $\sum_{n=0}^{\infty} x^n = 1/(1-x)$. For $x > 1$, on the other hand, the sum does not converge. The grand partition function is therefore defined if and only if $\exp[-(\alpha+\beta\varepsilon_r)] < 1$, which implies $-\alpha < \beta\varepsilon_r$. This condition must be valid for all single-particle states r, and since $\varepsilon_0 = 0$ is the smallest energy eigenvalue, we conclude that $\mathcal{Z}$ is defined if and only if $-\alpha \equiv \mu/kT < 0$. This means that *the chemical potential of an ideal boson gas is necessarily negative.*

Evaluating the sums, we find

$$\mathcal{Z} = \frac{1}{1 - e^{-(\alpha+\beta\varepsilon_0)}} \frac{1}{1 - e^{-(\alpha+\beta\varepsilon_1)}} \frac{1}{1 - e^{-(\alpha+\beta\varepsilon_2)}} \times \cdots,$$

or

$$\ln \mathcal{Z} = -\ln\left[1 - e^{-(\alpha+\beta\varepsilon_0)}\right] - \ln\left[1 - e^{-(\alpha+\beta\varepsilon_1)}\right] - \ln\left[1 - e^{-(\alpha+\beta\varepsilon_2)}\right] + \cdots.$$

In other words,

$$\ln \mathcal{Z} = -\sum_r \ln\left[1 - e^{-(\alpha+\beta\varepsilon_r)}\right].$$

We will express $\ln \mathcal{Z}$ in a tidier form in a moment.

6.2.2   *Mean occupation numbers*

As Eq. (6.3) shows, the mean occupation number of the single-particle state r^* is obtained by differentiating $\ln \mathcal{Z}$ with respect to the corresponding eigenvalue ε_{r^*}. From our previous expression for $\ln \mathcal{Z}$, we see that the only term involving this eigenvalue is

$$-\ln\left[1 - e^{-(\alpha+\beta\varepsilon_{r^*})}\right],$$

and differentiating yields

$$-\frac{\beta e^{-(\alpha+\beta\varepsilon_{r^*})}}{1 - e^{-(\alpha+\beta\varepsilon_{r^*})}}.$$

According to Eq. (6.3), we must now multiply by $-1/\beta$. Rearranging slightly, we arrive at the following result: For an ideal gas of bosons, the mean occupation number of the single-particle state r is given by

$$\boxed{\bar{n}_r = \frac{1}{e^{\alpha+\beta\varepsilon_r} - 1}}. \tag{6.7}$$

This distribution of mean occupation numbers is called the *Bose-Einstein* distribution, and bosons are therefore said to obey the Bose-Einstein statistics. Contrary to our previous results which were quite formal, Eq. (6.7) provides a very concrete relation between $\bar{n}_r$ and ε_r, for given values of α and β.

Equation (6.7) can be inverted to give

$$e^{-(\alpha+\beta\varepsilon_r)} = \frac{\bar{n}_r}{1 + \bar{n}_r}.$$

This means that we can express the logarithm of the grand partition function as

$$\boxed{\ln \mathcal{Z} = \sum_r \ln(1 + \bar{n}_r)}. \tag{6.8}$$

This proves a statement made earlier: the grand partition function can be expressed in terms of the mean occupation numbers, as can the mean number of particles, the mean energy, and the pressure.

6.3   Fermi-Dirac statistics

6.3.1   Grand partition function

We now turn to the case of an ideal gas consisting of fermion particles. Our first task is to evaluate the grand partition function, which is given by an infinite product of factors of the form

$$\sum_{n_r} e^{-(\alpha+\beta\varepsilon_r)n_r} = \sum_{n_r=0}^{1} \left[e^{-(\alpha+\beta\varepsilon_r)} \right]^{n_r},$$

where we have indicated that for fermions, the occupation numbers can take the values 0 and 1 only. This sum is trivial, and we find that the grand partition function is given by

$$\mathcal{Z} = \left[1 + e^{-(\alpha+\beta\varepsilon_0)} \right] \left[1 + e^{-(\alpha+\beta\varepsilon_1)} \right] \left[1 + e^{-(\alpha+\beta\varepsilon_2)} \right] \times \cdots$$

or

$$\ln \mathcal{Z} = \ln\left[1 + e^{-(\alpha+\beta\varepsilon_0)} \right] + \ln\left[1 + e^{-(\alpha+\beta\varepsilon_1)} \right] + \ln\left[1 + e^{-(\alpha+\beta\varepsilon_2)} \right] + \cdots.$$

In other words,

$$\ln \mathcal{Z} = \sum_{r} \ln\left[1 + e^{-(\alpha+\beta\varepsilon_r)} \right].$$

We will rewrite this in terms of the mean occupation numbers in a moment.

6.3.2   Mean occupation numbers

Calculation of the mean occupation numbers proceeds exactly as in the previous section, starting from Eq. (6.3). The end result is the following: For an ideal gas of fermions, the mean occupation number of the single-particle state r is given by

$$\boxed{\bar{n}_r = \frac{1}{e^{\alpha+\beta\varepsilon_r} + 1}} . \tag{6.9}$$

This distribution of mean occupation numbers is called the *Fermi-Dirac* distribution, and fermions are therefore said to obey the Fermi-Dirac statistics. Again, Eq. (6.9) provides an explicit relation between $\bar{n}_r$ and ε_r, for given values of α and β.

Equation (6.9) can be inverted to give

$$e^{-(\alpha+\beta\varepsilon_r)} = \frac{\bar{n}_r}{1 - \bar{n}_r}.$$

This means that we can express the logarithm of the grand partition function as

$$\boxed{\ln \mathcal{Z} = -\sum_{r} \ln(1 - \bar{n}_r)} . \tag{6.10}$$

It is worth spending of moment to appreciate the similarities and differences between Eqs. (6.7) and (6.9), and between Eqs. (6.8) and (6.10).

6.4   Summary

We have now completed our reformulation of the statistical mechanics of ideal gases, and it is time to summarize the major ideas and results.

In this formulation, the focus shifts from the states of the entire system, which are difficult to characterize, to the single-particle states r, which can be fully characterized with just four quantum numbers. The link between these two very different sets of states is provided by the set of all occupation numbers, $\{n_r\}$. Indeed, because identical particles are indistinguishable in quantum mechanics, all that can be said about the state of the entire system is how many particles occupy each and everyone of the single-particle states; this information is provided by the occupation numbers. We recall that n_r is equal to the number of particles that occupy the single-particle state r.

Our ideal gas is not isolated. Instead, it interacts with a reservoir, in such a way that both energy and particles can be exchanged. (Our system therefore adopts the Gibbs distribution of probabilities.) Because of this interaction, the number of particles occupying a specific state r does not stay constant, but keeps on fluctuating. We are therefore lead to consider the *average value* of the occupation numbers, which is given by

$$\bar{n}_r = \frac{1}{e^{\alpha + \beta \varepsilon_r} \mp 1} , \tag{6.11}$$

where the *upper sign* applies to bosons, while the *lower sign* applies to fermions. Here, α and β are parameters related to the temperature T and the chemical potential μ of the reservoir: $\alpha = -\mu/kT$ and $\beta = 1/kT$. As we have shown, α is always positive for bosons, but there is no such restriction for fermions.

From the mean occupation numbers we may calculate all thermodynamic quantities of interest. Using the same sign convention as above, the logarithm of the grand partition function is given by

$$\ln \mathcal{Z} = \pm \sum_r \ln(1 \pm \bar{n}_r) . \tag{6.12}$$

The mean total number of particles is

$$N = \sum_r \bar{n}_r , \tag{6.13}$$

and the mean total energy is given by

$$E = \sum_r \bar{n}_r \, \varepsilon_r . \tag{6.14}$$

In these expressions, the sums are over all the single-particle states r, and ε_r denotes the corresponding energy eigenvalues. It is precisely this property — that the sums are over the single-particle states and not the states of the entire system — that makes this reformulation so useful.

The preceding four equations are the fundamental equations of quantum statistics. From these we can derive the following expression for the entropy:

$$S = k \sum_r \left[\pm (1 \pm \bar{n}_r) \ln(1 \pm \bar{n}_r) - \bar{n}_r \ln \bar{n}_r \right] . \tag{6.15}$$

This derivation is the subject of a homework problem. We also recall that the pressure can be obtained directly from the general relation

$$\boxed{PV = \frac{2}{3} E}, \tag{6.16}$$

which holds both for bosons and fermions.

The present formulation of quantum statistics therefore has the virtue of being very economical. Its other important virtue, that it is easy to use in practical computations, will reveal itself in the following sections.

6.5   Classical limit

As our first application of the methods of quantum statistics, it is perhaps a good idea to make sure that we can at least reproduce the familiar results pertaining to the classical limit. After taking this tentative first step, we will be in a good position to generalize our discussion to more interesting situations.

6.5.1   Classical approximation

We must first specify what we mean by the classical limit. The classical picture of an ideal gas is that of little billiard balls moving around in a box. In this classical picture, there is no difficulty of principle in specifying the position of each particle with arbitrary accuracy. The quantum picture, of course, is drastically different. Here, the position of each particle cannot be specified with arbitrary accuracy, because of the Heisenberg uncertainty principle. The best we can do is to picture our particles as fuzzy wave packets moving around in the box. The uncertainty Δx in the position of a given wave packet is related to Δp, the uncertainty in its momentum, by the Heisenberg relation

$$\Delta x \, \Delta p \sim \hbar.$$

The fuzzy nature of the gas particles never disappears completely, even in the classical limit. What happens, however, is that as we approach the classical description, the wave packets become increasingly localized. In other words, the wave-packet extension Δx becomes very small compared with the mean distance between gas particles. Because the mean volume per particle is V/N, we have that the condition

$$\Delta x \ll (V/N)^{1/3}$$

describes the classical limit.

We now want to convert this condition into something more useful. For this purpose, we use the Heisenberg relation to convert Δx into $\hbar/\Delta p$. Now, Δp is at most of the order of p itself, which is approximately equal to $\sqrt{2m(E/N)}$, where E/N is the mean energy per particle, which is of the order of kT. Gathering the results, and discarding meaningless numerical coefficients, we obtain

$$\Delta x \leq \frac{\hbar}{\sqrt{mkT}}.$$

Substituting this into the preceding equation, we find that the classical limit corresponds to the mathematical statement

$$\lambda^3 \ll V/N,$$

where $\lambda \equiv \Delta x$ is the length scale associated with the fuzziness of our particles:

$$\boxed{\lambda = \frac{2\pi\hbar}{\sqrt{2\pi m k T}}} \;. \tag{6.17}$$

For convenience, numerical factors have been inserted in this expression for λ.

We now express the condition $\lambda^3 \ll V/N$ in a different and more useful form. For this purpose, we recall from a previous homework problem that the chemical potential of a classical gas is given by

$$\mu = -kT \ln\left[\frac{V/N}{\lambda^3}\right].$$

This tells us immediately that the condition for the classical limit can be written as

$$\boxed{e^\alpha \gg 1 \qquad \leftrightarrow \qquad \text{classical limit}} \;. \tag{6.18}$$

This is the form that we will use in our calculations. In fact, we know already that this condition is true: a helium gas at atmospheric pressure and at room temperature has $\alpha = -\mu/kT \simeq 13$, which yields $e^\alpha \simeq 10^5$. This is indeed much larger than 1!

Notice that we have cheated a little. We have used a previously derived expression for the chemical potential to give a precise formulation of the classical approximation, which we will now use to calculate the thermodynamic quantities of a classical gas, including the chemical potential! This is not as bad as it seems. We may now forget that the chemical potential is given by the expression above. All we shall need is Eq. (6.18), which merely tells us that in our calculations, we can treat e^α as a very large quantity. This is *far less* information than the detailed knowledge of how μ varies with temperature T and number density N/V.

6.5.2   *Evaluation of α*

What better proof of this statement than a rederivation our previous expression for the chemical potential, directly from the equations listed in Sec. 4?

In the classical limit, where e^α, and therefore also $e^{\alpha + \beta \varepsilon_r}$, are much larger than one, Eq. (6.11) reduces to

$$\boxed{\bar{n}_r = e^{-(\alpha + \beta \varepsilon_r)}} \;, \tag{6.19}$$

which implies that the mean occupation numbers are all very small. This means that in the classical limit, the particles are distributed over an extremely large number of single-particle states. Equation (6.19) also shows that in the classical limit, there is no distinction between bosons and fermions.

We now use Eq. (6.19) to calculate N, the mean number of particles within the gas:

$$N = \sum_r \bar{n}_r = e^{-\alpha} \sum_r e^{-\beta \varepsilon_r}.$$

By inverting this relation, we will be able to express the chemical potential $\mu = -kT\alpha$ in terms N, which is exactly what we want. But we must first evaluate the sum $\sum_r e^{-\beta \varepsilon_r}$, which is nothing but the partition function of a *single* particle confined to a box of volume V.

To evaluate this, we first observe that a number $\Omega(\varepsilon_r) \neq 1$ of single-particle states r will share the same energy eigenvalue ε_r. The quantity $\Omega(\varepsilon_r)$ is of course the statistical weight for a single particle in a box. The sum over single-particle states can therefore be written as

$$\sum_r = \sum_{\varepsilon_r} \Omega(\varepsilon_r),$$

where we are now summing over the *distinct* energy eigenvalues. Because these eigenvalues are quantized, this is a discrete sum. But because the *vast* majority of the energy eigenvalues are so extremely close to each other, we may in fact treat ε_r as a continuous variable. We shall therefore replace the discrete sum by an integral,

$$\boxed{\sum_r \to \int_0^\infty \omega(\varepsilon)\,d\varepsilon}\,, \tag{6.20}$$

by re-introducing the *density of states* $\omega(\varepsilon)$. This, we recall, is defined so that $\omega(\varepsilon)\,d\varepsilon$ is the number of single-particle states with energy eigenvalues in the interval between ε and $\varepsilon + d\varepsilon$. The density of states has already been calculated in Sec. B 3:

$$\boxed{\omega(\varepsilon) = g\,\frac{2\pi V}{(2\pi\hbar)^3}\,(2m)^{3/2}\,\varepsilon^{1/2} = \frac{2gV}{\sqrt{\pi}\lambda^3}\,\beta^{3/2}\,\varepsilon^{1/2}}\,, \tag{6.21}$$

where Eq. (6.17) was used to obtain the second expression. The factor g is new. We need to insert it to account for the number of internal states associated with the particle's spin, something which we had neglected to do before. The number g is equal to the number of values that m_s can take between the bounds $-s$ and s. Clearly, this is

$$\boxed{g = 2s + 1}\,, \tag{6.22}$$

where s is the particle's spin.

   We are now ready to calculate the chemical potential. Gathering our previous results, we have

$$\begin{aligned} N &= e^{-\alpha}\int_0^\infty e^{-\beta\varepsilon}\omega(\varepsilon)\,d\varepsilon \\ &= \frac{2gV}{\sqrt{\pi}\lambda^3}\,e^{-\alpha}\,\beta^{3/2}\int_0^\infty e^{-\beta\varepsilon}\varepsilon^{1/2}\,d\varepsilon \\ &= \frac{2gV}{\sqrt{\pi}\lambda^3}\,e^{-\alpha}\int_0^\infty e^{-x}x^{1/2}\,dx, \end{aligned}$$

where, in the last line, we have put $x \equiv \beta\varepsilon$. We have encountered integrals of this sort before, and in fact, this integral is $\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi}$. We have therefore obtained the simple result $N = gVe^{-\alpha}/\lambda^3$, or

$$\boxed{e^{-\alpha} = \frac{N\lambda^3}{gV}}\,. \tag{6.23}$$

This means that $\mu = -kT\ln(gV/N\lambda^3)$, and apart from the new factor g, this is exactly what we had written in subsection a.

6.5.3   Other thermodynamic quantities

The same method can be used to calculate other quantities of interest. Let us start with the mean energy, which according to Eqs. (6.14), (6.19), and (6.20), is given by

$$E = \sum_r \bar{n}_r\,\varepsilon_r = e^{-\alpha}\sum_r \varepsilon_r\,e^{-\beta\varepsilon_r} = e^{-\alpha}\int_0^\infty \varepsilon e^{-\beta\varepsilon}\omega(\varepsilon)\,d\varepsilon.$$

Substituting Eqs. (6.21) and (6.23), this becomes

$$E = \frac{2N}{\sqrt{\pi}\beta}\int_0^\infty e^{-x}x^{3/2}\,dx,$$

where $x \equiv \beta\varepsilon$. The integral evaluates to $\Gamma(\frac{5}{2}) = \frac{3}{4}\sqrt{\pi}$, and we arrive at the familiar result

$$\boxed{E = \frac{3}{2}NkT} \, . \tag{6.24}$$

From this and Eq. (6.16) we immediately obtain the classical limit for the equation of state of an ideal gas:

$$\boxed{PV = NkT} \, . \tag{6.25}$$

Let us see if we can also compute the entropy. From Eq. (6.15) and the fact that the occupation numbers are all much smaller than unity — we use the approximation $\ln(1 + \epsilon) \simeq \epsilon$ when ϵ is small — we obtain

$$S = k \sum_r (\bar{n}_r - \bar{n}_r \ln \bar{n}_r).$$

But Eq. (6.19) also gives $\ln \bar{n}_r = -\alpha - \beta\varepsilon_r$, so the entropy becomes $S = k(N + \alpha N + \beta E) = Nk(\alpha + \frac{5}{2})$. Substituting Eq. (6.23), we arrive at

$$\boxed{S = Nk \ln\left(\frac{gV}{N\lambda^3}\right) + \frac{5Nk}{2}} \, . \tag{6.26}$$

This result differs from Eq. (2.9) only by the inclusion of the factor g.

6.5.4   Conclusion

We see that we are able to reproduce all of the familiar classical results. It should be appreciated that calculations were handled quite easily and efficiently in our current formulation of quantum statistics. The major advantage of this new formulation is that it will allow us to go beyond the classical limit. We will be able to predict how quantum effects affect such relations as $E = \frac{3}{2}NkT$ and $PV = NkT$, and this will bring to light interesting new phenomena.

It will be useful, for what comes next, to introduce a terminology to characterize the importance of quantum effects in an ideal gas. We shall term *nondegenerate* gases which behave essentially classically. The criterion for this was derived in subsection a: it is $\lambda \ll (V/N)^{1/3}$, or $e^\alpha \gg 1$. We shall term *slightly degenerate* gases for which quantum effects are still quite small, but not negligible. The criterion for this is $\lambda < (V/N)^{1/3}$, or $e^\alpha > 1$. Finally, we shall term *highly degenerate* gases for which quantum effects are extremely important. The criterion for this is $\lambda \geq (V/N)^{1/3}$, or $e^\alpha \leq 1$.

6.6   Slightly degenerate ideal gases

We now generalize the discussion of the preceding section to ideal gases for which quantum effects are becoming important. As was mentioned before, we refer to this situation as the slightly degenerate limit, in which e^α is still larger than one, but only moderately so. Thus, we want to compute the leading order corrections to the classical results derived in the preceding section.

6.6.1   Evaluation of α

We first recall that for bosons (upper sign) and fermions (lower sign), the mean occupation numbers are given by

$$\bar{n}_r = \frac{1}{e^{\alpha + \beta\varepsilon_r} \mp 1}.$$

This is an exact result. But if e^α is fairly large, then the mean occupation numbers are given approximately by

$$\boxed{\bar{n}_r = e^{-(\alpha+\beta\varepsilon_r)} \pm e^{-2(\alpha+\beta\varepsilon_r)} + \cdots} \,, \tag{6.27}$$

where the dots represent terms of order $e^{-3\alpha}$ and smaller. These will be neglected throughout the calculation. We will first use Eq. (6.27) to express $e^{-\alpha}$ in terms of N, the mean number of particles.

We start with the relation $N = \sum_r \bar{n}_r$, in which we substitute Eq. (6.27). After converting the discrete sums into integrals, we obtain

$$N = e^{-\alpha} \int_0^\infty e^{-\beta\varepsilon}\omega(\varepsilon)\,d\varepsilon \pm e^{-2\alpha} \int_0^\infty e^{-2\beta\varepsilon}\omega(\varepsilon)\,d\varepsilon + \cdots.$$

We already know from Sec. 5 that the first integral gives gV/λ^3. To evaluate the second integral, we substitute Eq. (6.21) and find

$$\int_0^\infty e^{-2\beta\varepsilon}\omega(\varepsilon)\,d\varepsilon = \frac{2gV}{\sqrt{\pi}\lambda^3}\,\beta^{3/2}(2\beta)^{-3/2}\int_0^\infty e^{-x}x^{1/2}\,dx,$$

where $x \equiv 2\beta\varepsilon$. The integral gives $\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi}$, and gathering the results, we arrive at

$$N = \frac{gV}{\lambda^3}\,e^{-\alpha}\big(1 \pm 2^{-3/2}e^{-\alpha} + \cdots\big).$$

It is not difficult to invert this relation between N and $e^{-\alpha}$, especially if we keep in mind that all terms of order $e^{-3\alpha}$ and smaller must be neglected. We obtain

$$\boxed{e^{-\alpha} = \frac{N\lambda^3}{gV}\bigg(1 \mp 2^{-3/2}\frac{N\lambda^3}{gV} + \cdots\bigg)}. \tag{6.28}$$

This gives the chemical potential of a slightly degenerate ideal gas. Equation (6.28) shows that for bosons, quantum effects produce a slight *decrease* in $e^{-\alpha}$ with respect to the classical result. For fermions, quantum effects produce a slight *increase*. The physical interpretation of this result will come more easily when we consider the other thermodynamic quantities.

6.6.2   Energy and pressure

The mean energy of the gas is calculated as in the previous section. After substituting Eq. (6.27) into $E = \sum_r \bar{n}_r\,\varepsilon_r$ and converting into integrals, we obtain

$$E = e^{-\alpha} \int_0^\infty \varepsilon e^{-\beta\varepsilon}\omega(\varepsilon)\,d\varepsilon \pm e^{-2\alpha} \int_0^\infty \varepsilon e^{-2\beta\varepsilon}\omega(\varepsilon)\,d\varepsilon + \cdots.$$

After using Eq. (6.21), the first integral gives $3gV/2\lambda^3\beta$ as before, while the second integral is

$$\int_0^\infty \varepsilon e^{-2\beta\varepsilon}\omega(\varepsilon)\,d\varepsilon = \frac{2gV}{\sqrt{\pi}\lambda^3}\,\beta^{3/2}(2\beta)^{-5/2}\int_0^\infty e^{-x}x^{3/2}\,dx,$$

where $x \equiv 2\beta$. The integral gives $\Gamma(\frac{5}{2}) = \frac{3}{4}\sqrt{\pi}$, and gathering the results, we obtain

$$E = \frac{3}{2}\,kT\,\frac{gV}{\lambda^3}e^{-\alpha}\big(1 \pm 2^{-5/2}e^{-\alpha} + \cdots\big).$$

Although this is a perfectly valid expression for the energy, we would prefer to express E in terms of N. Using Eq. (6.28) gives

$$E = \frac{3}{2}\,kT\,N\bigg(1 \mp 2^{-3/2}\frac{N\lambda^3}{gV} + \cdots\bigg)\bigg[1 \pm 2^{-5/2}\frac{N\lambda^3}{gV}\big(1 \mp \cdots\big)\bigg].$$

Recalling that we must neglect all correction terms beyond $N\lambda^3/gV$, this simplifies to

$$E = \frac{3}{2}NkT\left(1 \mp 2^{-5/2}\frac{N\lambda^3}{gV} + \cdots\right). \qquad (6.29)$$

This result shows that the classical relation $E = \frac{3}{2}NkT$ is modified when quantum effects are taken into account. Equation (6.29) further shows that the energy is slightly *decreased* for bosons, while it is slightly *increased* for fermions.

Equations (6.16) and (6.29) give the equation of state for a slightly degenerate ideal gas:

$$PV = NkT\left(1 \mp 2^{-5/2}\frac{N\lambda^3}{gV} + \cdots\right). \qquad (6.30)$$

This result is quite interesting, because it shows that the classical relation $PV = NkT$ is modified by the quantum nature of the gas. We see that the pressure is *lower* for bosons, and *higher* for fermions.

It is worth spending some time trying to understand this result. Why is the pressure higher for fermions? The answer comes from the fact that fermions do not like to occupy the same state, so that they will resist compression more vigourously than classical particles. The extra pressure provided by quantum effects can therefore be understood as being there to help enforce Pauli's exclusion principle. This effect is most dramatic in the highly degenerate limit, to which we turn in the next three sections.

Why is the pressure lower for bosons? This is not as easy to understand. Bosons, of course, do not obey an exclusion principle, and this explains why the pressure of a boson gas must be lower than that of a fermion gas. This, however, does not explain why the pressure must be lower than the classical result. The answer lies with the fact that unlike fermions and classical particles, bosons prefer to be close together, so that they will resist compression less vigourously than fermions and classical particles. This effect is most dramatic in the highly degenerate limit, which will be the topic of Secs. 10, 11, and 12.

6.6.3   Higher-order corrections

Higher-order corrections to the equation of state can be calculated in the same way, by simply keeping more terms in our expansions in powers of $e^{-\alpha}$, or in powers of $N\lambda^3/gV$. We will not go through this exercise, but it should at least be clear that the end result will be an expansion of the form

$$\begin{aligned} PV &= NkT\left[1 + a_1\left(\frac{N\lambda^3}{gV}\right) + a_2\left(\frac{N\lambda^3}{gV}\right)^2 + \cdots\right] \\ &= NkT\sum_{n=0}^{\infty} a_n\left(\frac{N\lambda^3}{gV}\right)^n. \end{aligned}$$

Such an expression for the equation of state of a gas is called a *virial expansion*, and the coefficients a_n are called *virial coefficients*. The first few are given by $a_0 = 1$, $a_1 = \mp 2^{-5/2} \simeq \mp 0.17678$, $a_2 \simeq -0.00330$, and $a_3 \simeq \mp 0.00011$. In this section we have calculated only the first-order correction. Wouldn't it be a wonderful challenge to calculate the next term?

6.7 Highly degenerate Fermi gas. I: Zero temperature

We have seen in the preceding section that quantum effects modify the behaviour of the thermodynamic quantities of an ideal gas, with respect to their classical behaviour. For example, the pressure of a slightly degenerate Fermi gas was found to be slightly larger than NkT/V, while it was found to be smaller than this for a Bose gas. In this and the following sections we will explore the depths of quantum statistics by going to the highly degenerate limit, where quantum effects truly dominate. By doing so we will encounter drastic deviations from the classical predictions, and interesting new phenomena will be discovered.

We recall from Sec. 5d that the highly degenerate limit corresponds to a situation in which the quantum wave packets representing the gas particles are no longer well separated from each other. (The opposite situation, in which the wave packets are well separated, is the classical limit.) We also recall from Sec. 5a that the typical extension of the wave packets is given by λ, as defined by Eq. (6.17). What is noteworthy here is that λ is inversely proportional to the square root of the temperature: $\lambda \propto T^{-1/2}$. The wave packets therefore increase in size as the temperature decreases, and there comes a point where λ becomes comparable with $(V/N)^{1/3}$, the mean distance between wave packets. At this point the gas can no longer be pictured in terms of individual particles: the wave packets all overlap with each other. In this situation, clearly, quantum effects play a very important role.

We shall begin our discussion of the highly degenerate limit by examining a fermion gas at zero temperature, where things are quite extreme.

6.7.1 Fermi gas at $T = 0$

It should be quite clear that at $T = 0$, a system of fermions will be found in its state of lowest energy, or ground state. This, however, does not mean that each particle will be found in a single-particle state of zero energy. Indeed, Pauli's exclusion principle prevents this from happening: Once the first g particles have been put in the single-particle states of zero energy, additional particles must be put in the excited states, until all N particles are accounted for. Obviously then, the state of the entire system will not be a state of zero energy, although it is still the state of lowest energy.

Following these considerations, it is easy to predict what values the occupation numbers will take at zero temperature. Because the first particle is placed in the state $r = 0$, we automatically have $n_0 = 1$. The second particle, on the other hand, is placed in the state $r = 1$, so we have $n_1 = 1$. Proceeding similarly for all N particles, we find that every occupation number up to n_{N-1} is equal to 1. Because we now have accounted for all the particles, all other occupation numbers must be zero. So we have

$$n_r = \begin{cases} 1 & r \leq N - 1 \\ 0 & r \geq N \end{cases} .$$

The occupation numbers behave as a step function. The energy of the last occupied single-particle state ($r = N - 1$) is called the *Fermi energy*, and is denoted ε_F.

This step-function behaviour should also be apparent in the mean occupation numbers, which for a Fermi gas are given by

$$\bar{n}_r = \frac{1}{e^{(\varepsilon_r - \mu)/kT} + 1}.$$

In the limit $T \to 0$, we have to pay close attention to the sign of $\varepsilon_r - \mu$. Suppose first that $\varepsilon_r < \mu$. In this case $e^{(\varepsilon_r - \mu)/kT}$ is an extremely *small* number, and we find
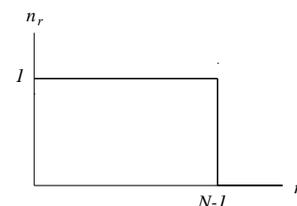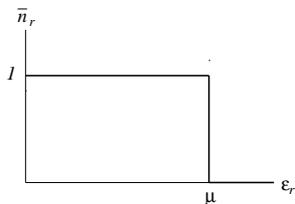
that $\bar{n}_r = 1$. Suppose now that $\varepsilon_r > \mu$. In this case $e^{(\varepsilon_r - \mu)/kT}$ is an extremely *large* number, and we find that $\bar{n}_r = 0$. Thus, the mean occupation numbers jump abruptly from 1 to 0 as ε_r reaches the value μ. This is clearly the same behaviour as what we have found for the occupation numbers themselves, and we may identify the Fermi energy with the chemical potential at zero temperature:

$$\boxed{\varepsilon_F = \lim_{T \to 0} \mu(T)}. \tag{6.31}$$

We then observe that $e^{-\alpha} = e^{\varepsilon_F/kT} \gg 1$ when $T \to 0$. Recall that the classical limit corresponds to the opposite situation, $e^{-\alpha} \ll 1$.

6.7.2   Evaluation of the Fermi energy

It is not difficult to express the Fermi energy in terms of other properties of the gas, most importantly, its density. We start from the basic expression for the mean number of particles,

$$N = \sum_r \frac{1}{e^{(\varepsilon_r - \mu)/kT} + 1},$$

which we convert into the integral

$$N = \int_0^\infty \frac{\omega(\varepsilon)\, d\varepsilon}{e^{(\varepsilon - \mu)/kT} + 1}.$$

Such an integral would be difficult to evaluate in general. But we know that the function $1/(e^{(\varepsilon - \mu)/kT} + 1)$ behaves as a step function in the limit $T \to 0$, so the integral becomes simply

$$N = \int_0^{\varepsilon_F} \omega(\varepsilon)\, d\varepsilon.$$

This is quite easy to evaluate if we substitute Eq. (6.21). A few lines of algebra yield

$$\boxed{\varepsilon_F = \frac{(2\pi\hbar)^2}{2m} \left(\frac{3N}{4\pi g V} \right)^{2/3}}. \tag{6.32}$$

This result shows that the Fermi energy increases when N/V, the density of the gas, increases. In the sequel we will find it useful to rewrite Eq. (6.21) in the form

$$\omega(\varepsilon) = \frac{3N}{2\varepsilon_F^{3/2}}\, \varepsilon^{1/2}.$$

6.7.3   Energy and pressure

The mean energy of the gas is calculated in a similar way. We shall leave this as a homework problem. You should find that the mean energy of a Fermi gas at zero temperature is equal to

$$\boxed{E = \frac{3}{5} N\varepsilon_F}. \tag{6.33}$$

From this it follows that the gas also has a pressure given by

$$\boxed{PV = \frac{2}{5} N\varepsilon_F}. \tag{6.34}$$

A Fermi gas at zero temperature has pressure! Clearly, this pressure is not created by the thermal agitation of the gas particles. This could not happen at $T = 0$. Instead, the pressure is kindly provided by professor Pauli: the fermions fight hard

not to occupy the same single-particle state, and it is this fight that is responsible for the pressure. In other words, Pauli's exclusion principle provides *all* the pressure in a Fermi gas at zero temperature. We shall see in Sec. 9 that this *degeneracy pressure* is put to good use in nature.

6.7.4   Entropy

Because our Fermi gas is at zero temperature, its entropy is automatically zero:

$$\boxed{S = 0} \ . \tag{6.35}$$

This explains why the chemical potential at zero temperature is the Fermi energy. Recall the fundamental definition of the chemical potential as an injection energy:

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S,V}.$$

In words, the chemical potential is equal to the change in energy when new particles are added to the system, keeping the entropy constant. Now, at zero temperature, a particle can be added to the system if and only if its energy is larger than, or equal to, the Fermi energy. If it is *larger* than ε_F, then the system's temperature will increase, and so also will its entropy. This does not correspond to the operational definition of the chemical potential. If, on the other hand, the particle's energy is *equal to* the Fermi energy, then the system will stay at $T = 0$, and its entropy will not increase from $S = 0$. This does correspond to the operational definition of the chemical potential. We may then conclude that the Fermi energy must be equal to the chemical potential at $T = 0$.

6.8   Highly degenerate Fermi gas. II: Low temperature

We now move away from $T = 0$, and ask how our previous results are affected by finite-temperature effects. For simplicity, we shall demand that kT be much smaller than the Fermi energy ε_F. By so doing, we will still be dealing with a highly degenerate Fermi gas.

6.8.1   Fermi-Dirac integrals

In the following calculations we will encounter integrals of the form

$$I[f] = \int_0^\infty \frac{f(\varepsilon) \, d\varepsilon}{e^{(\varepsilon - \mu)/kT} + 1} \ ,$$

where $f(\varepsilon)$ is some function. We have already recognized the fact that such integrals are difficult to evaluate in general, but that they simplify drastically when $T = 0$. When T is different from zero but small, $I[f]$ can be evaluated approximately. The result, which is derived in Sec. 9.17 of Reif's book, is

$$\boxed{I[f] = \int_0^\mu f(\varepsilon) \, d\varepsilon + \frac{\pi^2}{6} \, f'(\mu) \, (kT)^2 + \cdots} \ , \tag{6.36}$$

where $f' \equiv df/d\varepsilon$. The derivation of this result is not particularly difficult, but it is quite long and tedious. This is why we have chosen to omit it.

6.8.2 Chemical potential

We first compute the finite-temperature corrections to the chemical potential μ. We start once again with the following expression for the mean number of particles:

$$N = \int_0^\infty \frac{\omega(\varepsilon)\,d\varepsilon}{e^{(\varepsilon-\mu)/kT}+1} = \frac{3N}{2\varepsilon_F^{3/2}} \int_0^\infty \frac{\varepsilon^{1/2}\,d\varepsilon}{e^{(\varepsilon-\mu)/kT}+1}.$$

The integral is $I[\varepsilon^{1/2}]$, and substitution of Eq. (6.36) yields

$$N = \frac{3N}{2\varepsilon_F^{3/2}} \left[\frac{2}{3}\mu^{3/2} + \frac{\pi^2}{12}\mu^{-1/2}(kT)^2 + \cdots \right].$$

A little cleaning up then gives

$$\varepsilon_F^{3/2} = \mu^{3/2}\left[1 + \frac{\pi^2}{8}\left(\frac{kT}{\mu}\right)^2 + \cdots \right],$$

or

$$\varepsilon_F = \mu\left[1 + \frac{\pi^2}{12}\left(\frac{kT}{\mu}\right)^2 + \cdots \right].$$

Finally, by inverting this series we obtain the desired expression for the chemical potential:

$$\mu(T) = \varepsilon_F\left[1 - \frac{\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2 + \cdots \right]. \tag{6.37}$$

We see that finite-temperature effects decrease the chemical potential from its value at $T = 0$.

6.8.3 Energy and pressure

The mean energy is calculated in a similar way. We start from

$$E = \int_0^\infty \frac{\varepsilon\,\omega(\varepsilon)\,d\varepsilon}{e^{(\varepsilon-\mu)/kT}+1} = \frac{3N}{2\varepsilon_F^{3/2}} \int_0^\infty \frac{\varepsilon^{3/2}\,d\varepsilon}{e^{(\varepsilon-\mu)/kT}+1},$$

and we identify the second integral as $I[\varepsilon^{3/2}]$. Proceeding as before, we obtain

$$E = \frac{3N}{5\varepsilon_F^{3/2}}\mu^{5/2}\left[1 + \frac{5\pi^2}{8}\left(\frac{kT}{\mu}\right)^2 + \cdots \right].$$

If we now substitute Eq. (6.37) raised to the 5/2 power, this becomes

$$E = \frac{3N}{5\varepsilon_F^{3/2}}\varepsilon_F^{5/2}\left[1 - \frac{5\pi^2}{24}\left(\frac{kT}{\varepsilon_F}\right)^2 + \cdots \right]\left[1 + \frac{5\pi^2}{8}\left(\frac{kT}{\varepsilon_F}\right)^2 + \cdots \right].$$

After multiplication, we obtain

$$E = \frac{3}{5}N\varepsilon_F\left[1 + \frac{5\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2 + \cdots \right]. \tag{6.38}$$

We see that finite-temperature effects increase the mean energy of the gas. Not a surprising result!

From Eq. (6.38) we immediately obtain the finite-temperature correction to the pressure,

$$PV = \frac{2}{5}N\varepsilon_F\left[1 + \frac{5\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2 + \cdots \right], \tag{6.39}$$

which is also increased with respect to the zero-temperature result.

6.8.4 Heat capacity

We conclude this section by computing C_V, the heat capacity at constant volume. This is defined by $(đQ/dT)_V$, which is clearly equal to $\partial E/\partial T$. Differentiating Eq. (6.38) with respect to T yields

$$\boxed{C_V = \frac{\pi^2}{2} Nk \left(\frac{kT}{\varepsilon_F}\right)}. \tag{6.40}$$

The heat capacity of a Fermi gas is therefore proportional to T in the highly degenerate limit $kT \ll \varepsilon_F$. At high temperatures, on the other hand, the gas behaves classically, and the heat capacity saturates at the constant value

$$C_V = \frac{3}{2} Nk.$$
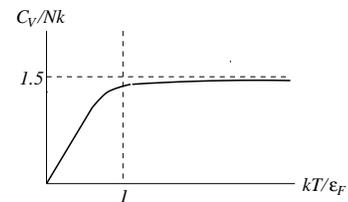
6.9 Highly degenerate Fermi gas. III: Applications

Highly degenerate Fermi gases occur quite commonly in nature. We describe a few examples in this section.

6.9.1 Conduction electrons in metals

Consider a typical metal, copper. When copper atoms combine to form a solid, each atom loses one or two electrons. The lost electrons move (more or less) freely within the metal, and therefore form a gas. It is this electron gas which is responsible for the thermal and electrical properties of the metal (heat conduction, current conduction, etc.).

We may model the electron gas as an ideal Fermi gas, thereby neglecting all interactions with the copper ions, and all mutual interactions. We first calculate the *Fermi temperature* of this gas,

$$T_F \equiv \frac{\varepsilon_F}{k} = \frac{(2\pi\hbar)^2}{2km} \left(\frac{3N}{4\pi g V}\right)^{2/3}.$$

Since we are dealing with an electron gas, $g = 2$ and $m = 9.1 \times 10^{-28}$ g. The density of copper is $N/V = 8.4 \times 10^{22}$ electrons/cm^3. This gives

$$T_F \simeq 80\ 000 \text{ K},$$

which is much higher than room temperature! In fact, the Fermi temperature is also much higher than copper's melting point, which is approximately 1 000 K. The conclusion is therefore that *conduction electrons in metals form a highly degenerate Fermi gas.*

The heat capacity of metals has long been measured to grow *linearly* with temperature. This experimental fact was at first very puzzling, because before the advent of quantum mechanics, C_V was expected to behave classically, $C_V = \frac{3}{2}Nk$. This would indeed be the correct behaviour if *all* the conduction electrons were contributing to the thermal properties of the metal. But this is not the case: Only those electrons which lie above the Fermi level actually contribute, and this constitutes only a very small fraction of all the electrons. This fraction is of the order of kT/ε_F, and therefore,

$$C_V \sim Nk \frac{kT}{\varepsilon_F}.$$

Apart from the numerical factor of $\pi^2/2$, this is exactly the result derived in the preceding section.

6.9.2   White dwarfs

During their normal lives, stars are supported against gravity by radiation pressure, which is provided by the nuclear reactions occurring in the core. When these reactions cease, the pressure drops, and the core collapses. (The outer layers are then blown away and form what is known as a planetary nebula.) If the core mass is less than a few solar masses, then the endpoint of the collapse is a *white dwarf.*

The density inside a white dwarf is so high that the electrons form a highly degenerate Fermi gas. (Because the electrons move with relativistic speeds, the laws of relativistic statistical mechanics must be used to work out the details. This lies somewhat beyond the scope of this course.) It is now the degeneracy pressure, provided by Pauli's exclusion principle, which supports the dead star against gravity.

The first white dwarf ever discovered was Sirius B, an invisible companion to the normal star Sirius A. Sirius B has a mass approximately equal to the solar mass, but its radius is comparable to that of the Earth: it is only 4 800 km.

6.9.3   Neutron stars

White dwarfs can exist only if their mass is less than the *Chandrasekhar limit* of 1.4 solar masses. For masses larger that this, the density becomes so high that inverse beta decay occurs,

$$p + e \to n + \nu;$$

the white dwarf becomes neutron rich. Because the lost electrons are no longer there to provide pressure, the dwarf collapses further to form a *neutron star.*

Neutron stars are supported against gravity by the neutrons, which also form a highly degenerate Fermi gas. A typical neutron star has a mass comparable to that of the sun, but its radius is of the order of 15 km.

Neutron stars can exist only if their mass is less than approximately 3 solar masses. For larger masses, the neutron degeneracy pressure is no longer sufficient to support the star against its own weight. The neutron star then collapses completely, and becomes a *black hole.*

6.10   Highly degenerate Bose gas. I: Zero temperature

Because bosons are not bound by Pauli's exclusion principle, a highly degenerate boson gas behaves very differently from a highly degenerate fermion gas. This is the topic of this and the following sections, which will conclude our discussion of the quantum statistics of ideal gases.

A Bose gas behaves quite simply at zero temperature: because there is no obstacle in doing so, all N particles just occupy the single-particle zero-energy states. (Because of spin degeneracy, there is a number g of zero-energy states.) Recalling from Eq. (6.11) that the mean occupation numbers of a Bose gas are given by

$$\bar{n}_r = \frac{1}{e^{\alpha + \beta \varepsilon_r} - 1},$$

where $\alpha = -\mu/kT$ and $\beta = 1/kT$, and using the fact that only the states for which $\varepsilon_r = 0$ are occupied at $T = 0$, we have that

$$N = \frac{1}{e^{\alpha} - 1}$$

at zero temperature. This can be inverted to give

$$\boxed{e^{\alpha} = 1 + \frac{1}{N} \simeq 1}. \tag{6.41}$$

Thus, the highly degenerate limit of a Bose gas is characterized by the fact that $\alpha \equiv -\mu/kT$ is a very small number, approximately equal to $1/N$.

It is obvious that if all the particles occupy the zero-energy states, then $E = 0$ at zero temperature. This immediately implies that $P = 0$, and $T = 0$ also implies $S = 0$. Not a lot happening in a Bose gas at $T = 0$!

6.11   Highly degenerate Bose gas. II: Low temperatures

The relation $e^{\alpha} \simeq 1$ stays valid at low temperatures, and we will use it to derive the following results. We are mainly interested in two things: How many particles stay in the zero-energy states when $T \neq 0$, and how does the system's mean energy increase with increasing temperature. To simplify the terminology, in the following we will refer to the zero-energy states as the single-particle *ground state*.

6.11.1   Bose-Einstein integrals

Before we begin, we must quote the following mathematical result:

$$\boxed{I_{\nu} \equiv \int_0^{\infty} \frac{x^{\nu} \, dx}{e^x - 1} = \Gamma(\nu+1)\,\zeta(\nu+1)} . \tag{6.42}$$

Here ν stands for any number, and $\Gamma(z)$ is our old friend the Gamma function. The function $\zeta(z)$ is new. It is called the *Riemann zeta function*, and it is defined by

$$\boxed{\zeta(z) = \sum_{k=1}^{\infty} \frac{1}{k^z}} . \tag{6.43}$$

By directly evaluating the sum, which converges quite rapidly, it is easy to show that

$$\zeta(\tfrac{3}{2}) \simeq 2.612, \qquad \zeta(\tfrac{5}{2}) \simeq 1.341.$$

The integrals $I_{1/2} = \frac{1}{2}\sqrt{\pi}\,\zeta(\tfrac{3}{2})$ and $I_{3/2} = \frac{3}{4}\sqrt{\pi}\,\zeta(\tfrac{5}{2})$ will be required in the following calculations.

6.11.2   Number of particles in ground state

We want to calculate N_0, the mean number of particles that stay in the single-particle ground state when $T \neq 0$. To do so, we go back to the fundamental relation

$$N = \sum_r \bar{n}_r = N_0 + \sum_{\text{excited states}} \bar{n}_r.$$

As usual, we replace the sum by an integral, which gives

$$N = N_0 + \int_0^{\infty} \frac{\omega(\varepsilon)\,d\varepsilon}{e^{\beta\varepsilon} - 1},$$

where we have used the approximation $e^{\alpha} = 1$. It may appear objectionable to the set the lower limit of integration to zero, as we are supposed to sum over the excited states only. However, because $\omega(\varepsilon) \propto \varepsilon^{1/2}$, the lower limit actually contributes nothing to the integral, and we make no significant error in setting it to zero.

Substituting Eq. (6.21) for the density of states gives

$$N_0 = N - \frac{2\pi g V}{(2\pi\hbar)^3} \, (2m)^{3/2} \, \beta^{-3/2} \int_0^\infty \frac{x^{1/2} \, dx}{e^x - 1} \, ,$$

where $x \equiv \beta\varepsilon$. The integral is just $I_{1/2}$, and we obtain

$$N_0 = N - \frac{2\pi g V}{(2\pi\hbar)^3} \, (2mkT)^{3/2} \times \frac{1}{2} \, \sqrt{\pi} \, \zeta(\tfrac{3}{2}).$$

Cleaning this up, we arrive at the desired result

$$\boxed{N_0 = N\left[1 - \left(\frac{T}{T_c}\right)^{3/2}\right], \qquad T \le T_c} \, , \tag{6.44}$$

where the constant T_c, called the *condensation temperature*, is given by

$$\boxed{T_c = \frac{(2\pi\hbar)^2}{2\pi m k} \left[\frac{N}{g\zeta(\tfrac{3}{2})V}\right]^{2/3}} \, . \tag{6.45}$$

Equation (6.44) tells us that as the temperature increases, the mean number of particles in the ground state decreases, until the condensation temperature T_c is reached, at which point all the particles occupy the excited states. This result certainly makes sense physically, as we expect that increasing the temperature will give some energy to the particles. What is remarkable, however, is that until $T = T_c$, there is still a *macroscopic* number of particles that stay in the ground state. In effect, the thermal energy is not shared by *all* the particles, but by only a fraction of them. As was indicated, Eq. (6.44) is valid only for $T \le T_c$. Beyond this point, the approximation $e^\alpha \simeq 1$ starts breaking down, and this prevents N_0 from going negative, a nonsensical result.

6.11.3 Energy and pressure

We now turn to the evaluation of the mean energy. We start with the integral

$$E = \int_0^\infty \frac{\varepsilon \, \omega(\varepsilon) \, d\varepsilon}{e^{\beta\varepsilon} - 1} \, ,$$

in which we substitute Eq. (6.21). After using Eq. (6.45) to simplify the resulting expression, we obtain

$$E = \frac{2N}{\sqrt{\pi}\zeta(\tfrac{3}{2})} \, \frac{1}{(kT_c)^{3/2}} \, \beta^{-5/2} \int_0^\infty \frac{x^{3/2} \, dx}{e^x - 1} \, ,$$

where $x \equiv \beta\varepsilon$. The integral is just $I_{3/2}$, and we arrive at

$$\boxed{E = \frac{3\zeta(\tfrac{5}{2})}{2\zeta(\tfrac{3}{2})} \, NkT \left(\frac{T}{T_c}\right)^{3/2}, \qquad T \le T_c} \, . \tag{6.46}$$

This is clearly very different from the classical result, $E = \frac{3}{2}NkT$, and also the Fermi-Dirac result (6.38).

From Eqs. (6.16) and (6.46) we immediately obtain an expression for the pressure,

$$\boxed{PV = \frac{\zeta(\tfrac{5}{2})}{\zeta(\tfrac{3}{2})} \, NkT \left(\frac{T}{T_c}\right)^{3/2}, \qquad T \le T_c} \, . \tag{6.47}$$

Again, this is markedly different from the classical equation of state, $PV = NkT$, and also the Fermi-Dirac relation (6.39). Also interesting is the fact that because $T_c \propto V^{-2/3}$, the right-hand side of Eq. (6.47) is actually proportional to V. This means that the pressure of a highly degenerate Bose gas is *independent* of volume!

6.11.4   Heat capacity

Finally, we evaluate C_V, the heat capacity at constant volume, by differentiating Eq. (6.46) with respect to T. This gives

$$\boxed{C_V = \frac{15\zeta(\frac{5}{2})}{4\zeta(\frac{3}{2})}\, Nk \left(\frac{T}{T_c}\right)^{3/2}, \qquad T \leq T_c}. \tag{6.48}$$

The numerical factor is approximately equal to 1.9. The heat capacity is therefore proportional to $T^{3/2}$ at low temperatures. At high temperatures, it settles to the classical value $\frac{3}{2}Nk$.

It can be shown that although the heat capacity is continuous at $T = T_c$, its slope $\partial C_V/\partial T$ is not. This discontinuity of the first derivative signals the occurrence of a *phase transition* at $T = T_c$: An ideal Bose gas cooled below the condensation temperature goes into a state in which a *macroscopic* number of particles have precisely zero energy. This transition is known as *Bose-Einstein condensation*, and the term "condensate" refers to that part of the system which occupies the single-particle ground state.
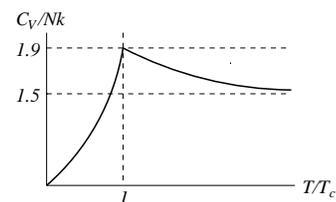
6.12   Highly degenerate Bose gas. III: Bose-Einstein condensation

Bose-Einstein condensation is manifested in various ways in nature. We describe a few examples in this section.

6.12.1   Superfluidity

Unless the pressure is extremely high, Helium 4 does not freeze when cooled toward $T = 0$. Instead, it undergoes a phase transition — the lambda point — at $T = T_\lambda = 2.17$ K. Below T_λ, the liquid becomes a *superfluid*, a liquid with no viscosity at all.

This phase transition bears a close resemblance to Bose-Einstein condensation. First, Helium 4 atoms are bosons, having zero spin. Second, the lambda-point temperature is very close to the condensation temperature,

$$T_c = \frac{(2\pi\hbar)^2}{2\pi mk}\left[\frac{N}{g\zeta\left(\frac{3}{2}\right)V}\right]^{2/3};$$

with $g = 1$, $m = 6.65 \times 10^{-24}$ g, and $V/N = 27.6$ cm^3/mole, this yields $T_c \simeq 3.1$ K.

However, there are significant differences between Bose-Einstein condensation and the lambda point. The most noticeable is in the behaviour of the heat capacity near the transition temperature: Whereas the heat capacity climbs to the finite value $1.9Nk$ in the case of Bose-Einstein condensation, it climbs to *infinity* in the case of the lambda point. (This λ-shaped behaviour of the heat capacity is at the origin of the term "lambda point".)

These differences can be attributed to the fact that unlike the ideal Bose gas, Helium 4 atoms have interactions that cannot be neglected. (Otherwise they would not form a liquid!) Nevertheless, the highly degenerate Bose gas provides a useful basis for our understanding of superfluidity.

6.12.2   Superconductivity

Some metals undergo a phase transition to a *superconducting* state as they are cooled below some critical temperature. This temperature is distinct for each metal, and ranges from approximately 1 K to 100 K. When the metal is in the superconducting state, it offers no resistance whatsoever to electric currents, and it expels any applied magnetic field (Meissner effect).

Again, in spite of differences due to interactions, the superconducting phase transition has strong similarities with Bose-Einstein condensation. But where are the bosons?

We already have mentioned that the thermal and electrical properties of metals are due to the conduction electrons, which are fermions. In the superconducting state, interactions within the metal allow electrons to form bound pairs, called Cooper pairs. These electron pairs obey Bose-Einstein statistics, and they are the bosons associated with the condensation phenomenon.

6.12.3   Holy Grail

To produce a near-ideal Bose gas and cool it down below T_c has long been considered a Holy Grail of experimental physics. This Grail was finally achieved in 1995: Bose-Einstein condensation was observed *directly* by an experimental group in Boulder, Colorado. The bosons are Rubidium 87 atoms. For the very low densities involved, $T_c \simeq 170$ nK.

Details on this remarkable discovery can be found in the August 1995 issue of *Physics Today*, or in the original paper by Anderson *et. al*, *Science* **269**, 198 (1995). Check out also the course web page: it displays, for three different temperatures, both below and above the condensation temperature, a beautiful graph of the velocity distribution of the Rubidium atoms

6.13   Problems

1. Show that the entropy of an ideal quantum gas is given by

$$S = k \sum_r \left[\pm (1 \pm \bar{n}_r) \ln(1 \pm \bar{n}_r) - \bar{n}_r \ln \bar{n}_r \right],$$

   with the upper sign applying to a boson gas, and the lower sign applying to a fermion gas.

2. In this problem we consider a physical system consisting of three particles, assumed to be bosons. Each of these can be in any one of three quantum states, with energy eigenvalues 0, ε, and 2ε, respectively.

   a) Using the rules of quantum statistics, calculate the partition function for this system. [Hint: First tabulate the states accessible to the entire system. These are specified by giving the values that each of the occupation numbers take, taking into account the fact that their sum must always be equal to three.]

   b) Evaluate the mean energy when the temperature is such that $\beta\varepsilon = 1$.

3. In this problem we consider the same system as above, but with an indefinite number of particles. (The system is interacting with a particle reservoir.)

   a) Calculate the grand partition function of this system.

   b) Evaluate the mean number of particles when the temperature is such that $\beta\varepsilon = 1$ and the chemical potential is such that $\alpha \simeq 0.3326$.

c) Evaluate the mean energy under the same conditions. Why doesn't your result agree with the answer you obtained in the previous problem?

4. Calculate the mean energy of an ideal Fermi gas at zero temperature. Express your result in terms of $N\varepsilon_F$. Calculate also the pressure of this gas.

5. (Reif 9.20) The atomic weight of sodium (Na) is 23, and the density of the metal is 0.95 g/cm^3. There is one conduction electron per atom.

   a) Calculate the Fermi temperature of the conduction elections.

   b) It is desired to cool a sample consisting of 100 cm^3 of Na metal from 1 K to 0.3 K. (At these low temperatures the lattice heat capacity of the metal is negligible compared to that due to to its conduction electrons.) The metal can be cooled by bringing it in thermal contact with liquid He^3 at 0.3 K. If 0.8 Joules of heat input are required to evaporate 1 cm^3 of liquid He^3, estimate how much of the liquid must be vaporized to cool the Na sample.

6. Show that the entropy of an extremely degenerate Fermi gas is given by

$$S = \frac{\pi^2}{2} Nk \frac{T}{T_F}.$$

   [Hint: Use the relation $E = TS - PV + \mu N$ derived in part A of this course.]

7. In this problem we consider a fictitious system consisting of N identical fermions, whose energy levels are given by

$$\varepsilon_r = r\,\varepsilon_1, \qquad r = 0, 1, 2, \ldots,$$

   where ε_1 is some unit of energy. It is assumed that $N \gg 1$, and that the system is in thermal equilibrium.

   a) Calculate $\omega(\varepsilon)$, the density of single-particle states for this system.

   b) Calculate ε_F, the system's Fermi energy.

   c) Assuming that the system is at zero temperature, calculate E, its mean energy.

   d) How high must the temperature be for E to differ significantly from the result of part c)? (Give a precise criterion.)

8. We have seen in the text that a typical metal contains electrons which move freely within the metal and therefore form an ideal Fermi gas. In certain metals, these electrons are confined to move in a two-dimensional plane. The single-electron states are therefore determined by the two-dimensional Schrödinger equation,

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\psi = E\psi.$$

   a) Show that the density of single-particle states for this two-dimensional system is given by

$$\omega(\varepsilon)\,d\varepsilon = \frac{4\pi mA}{(2\pi\hbar)^2}\,d\varepsilon,$$

   where A is the area of a two-dimensional cell within the metal.

   b) Calculate the system's mean energy at zero temperature, and express it in terms of N and ε_F.

CHAPTER 7
BLACK-BODY RADIATION

The term black-body radiation refers to a gas of photons in thermal equilibrium at a temperature T. Our purpose in this section is to apply the tools of statistical mechanics to work out the physical properties of such a gas. Our discussion relies heavily on the material presented in part F.

7.1   Photon statistics

Because they have a spin of unity, photons are in fact bosons, and as such they must obey the Bose-Einstein statistics. The mean occupation numbers for a gas of photons are therefore given by

$$\bar{n}_r = \frac{1}{e^{(\varepsilon_r - \mu)/kT} - 1}.$$

Here, r labels the single-photon states, and ε_r denotes their energy eigenvalues, which will be calculated in Sec. 2.

In fact, photons obey a special case of the Bose-Einstein statistics, which we shall call photon statistics. Indeed, it is easy to see that for photons, the chemical potential must vanish. This follows at once from the definition

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S,V},$$

which states that the chemical potential is equal to the change in the system's energy when photons are added *at constant entropy*. Since adding photons to a system is precisely what is meant by adding heat, and since adding heat is always accompanied by a change in entropy, adding photons can clearly *not* be done at constant entropy. In other words, the only amount of energy that can be added at constant entropy to a photon gas is zero. This, then, implies that the chemical potential must vanish:

$$\mu = 0.$$

The mean occupation numbers of a photon gas are therefore found to obey a special case of the Bose-Einstein distribution, given by

$$\boxed{\bar{n}_r = \frac{1}{e^{\varepsilon_r/kT} - 1}}. \tag{7.1}$$

This is called the *Planck* distribution.

7.2   Energy eigenvalues

Equation Eq. (7.1) is useless without an explicit expression for ε_r, the energy eigenvalues. We know that these are related to ω_r, the photon eigenfrequencies, by the Einstein relation,

$$\boxed{\varepsilon_r = \hbar\omega_r}\,.\tag{7.2}$$

We must therefore find a way to characterize the single-photon states r, and calculate their associated frequencies.

Whether it is treated as a classical quantity or as a quantum field operator, the electric field $\boldsymbol{E}$ always obeys the wave equation

$$\left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}+\nabla^2\right)\boldsymbol{E}=0,$$

where c is the speed of light. (The magnetic field $\boldsymbol{B}$ also satisfies this equation.) This is the equation we will use to find the single-photon states and their eigenfrequencies. In effect, the wave equation plays the same role for photons as the Schrodinger equation plays for massive particles.

By analogy with the solutions to the Schrodinger equation, we may look for solutions of the form

$$\boldsymbol{E}=\boldsymbol{\epsilon}\,\sin(k_x x)\,\sin(k_y y)\,\sin(k_z z)\,e^{-i\omega t},$$

where $\boldsymbol{\epsilon}$ is the field's polarization, and $\boldsymbol{k}$ is the wave vector. Substituting this into the wave equation indeed gives zero, provided that the frequency ω be given by

$$\omega^2=c^2\big(k_x{}^2+k_y{}^2+k_z{}^2\big).$$

This is the well-known *dispersion relation* for photons.

To find what k_x, k_y, and k_z are, we must examine the boundary conditions satisfied by $\boldsymbol{E}$. If we imagine that the radiation is contained in a box of volume V, then it is natural to impose that the field must vanish at the walls of this box. In effect, the box is imagined to be made of mirrors, and it is these mirrors that confine the radiation to the box. The condition that $\boldsymbol{E}$ vanishes at the walls simply means that the radiation is perfectly reflected by the mirrors. For simplicity, we suppose that the box is cubical, with length L and volume $V=L^3$. If we place two of the walls at $x=0$ and $x=L$, then the condition $\boldsymbol{E}=0$ at $x=L$ leads immediately to the equation $k_x L = n_x \pi$, where n_x is some integer. Proceeding similarly with the other dimensions gives $k_y L = n_y \pi$, $k_z L = n_z \pi$, and finally,

$$\boxed{\omega_r=\frac{\pi c}{V^{1/3}}\,\sqrt{n_x{}^2+n_y{}^2+n_z{}^2}}\,.\tag{7.3}$$

The single-photon states are therefore characterized by a set $r=\{n_x,n_y,n_z\}$ of three quantum numbers, and by the eigenfrequencies (7.3). These quantum numbers take integer values only. Furthermore, because changing the sign of a quantum number changes only the sign of $\boldsymbol{E}$, which does not produce a new quantum state, the quantum numbers may be assumed to take *positive* values only.

The characterization of single-photon states is therefore very similar to that of the states of massive particles: each requires a set of three positive integers. The energy eigenvalues, however, are different. While the energy is proportional to $n_x{}^2+n_y{}^2+n_z{}^2$ in the case of massive particles, it is proportional to its *square root* in the case of photons.

7.3   Density of states

Calculation of the thermodynamic quantities of a photon gas involves summing over the single-photon states. We are now used to the fact that these discrete sums can be converted into continuous integrals. This requires the introduction of the *density of states*, which we now calculate.

We denote the density of single-photon states by $\rho(\omega)$. This is defined by

$$\rho(\omega)\, d\omega \quad \equiv \quad \text{number of single-photon states in the frequency}$$
$$\text{interval between } \omega \text{ and } \omega + d\omega.$$

We have calculated similar things before, for example, in Sec. B.3.

Using Eq. (7.3), we see that $\rho(\omega)\, d\omega$ is equal to the number of points within the region of the fictitious n_x-n_y-n_z space defined by

$$\frac{V^{1/3}}{\pi c}\,\omega < \sqrt{n_x{}^2 + n_y{}^2 + n_z{}^2} < \frac{V^{1/3}}{\pi c}\,(\omega + d\omega).$$

Geometrically, this is equal to the volume of a spherical shell of radius $V^{1/3}\omega/\pi c$ and thickness $V^{1/3}d\omega/\pi c$. The result, however, must be multiplied by $1/8$, to account for the fact that the quantum numbers are all positive definite. Furthermore, the result must also be multiplied by 2, to account for the fact that photons have two polarization states. This factor of 2 is analogous to the factor g in Eq. (6.21), which accounts for the number of internal (spin) states of massive particles.

We therefore have

$$\rho(\omega)\, d\omega = 2\,\frac{1}{8}\,4\pi \left(\frac{V^{1/3}\omega}{\pi c}\right)^2 \frac{V^{1/3}d\omega}{\pi c},$$

or

$$\boxed{\rho(\omega)\, d\omega = \frac{V}{\pi^2 c^3}\,\omega^2\, d\omega}\,. \tag{7.4}$$

This is the density of single-photon states, which is used to convert discrete sums over these states into continuous integrals over ω.

7.4   Energy density

We are now ready to compute the thermodynamic quantities of the photon gas. Of particular interest is the radiation's mean energy, which is given by the usual relation

$$\begin{aligned}
E &= \sum_r \bar{n}_r\, \varepsilon_r = \sum_r \frac{\varepsilon_r}{e^{\varepsilon_r/kT} - 1} = \sum_r \frac{\hbar\omega_r}{e^{\hbar\omega_r/kT} - 1} \\
&= \int_0^\infty \frac{\hbar\omega\,\rho(\omega)\, d\omega}{e^{\hbar\omega/kT} - 1} = \frac{V\hbar}{\pi^2 c^3} \int_0^\infty \frac{\omega^3\, d\omega}{e^{\hbar\omega/kT} - 1},
\end{aligned}$$

where we have used Eqs. (7.2) and (7.4) in the second and fifth steps, respectively.

We shall denote E/V, the energy density, by the symbol u. We therefore have

$$u = \int_0^\infty I(\omega)\, d\omega,$$

where the function $I(\omega)$, called the spectral intensity, is given by

$$\boxed{I(\omega) = \frac{\hbar}{\pi^2 c^3}\,\frac{\omega^3}{e^{\hbar\omega/kT} - 1}}\,. \tag{7.5}$$

This result was first obtained by Planck in 1990, and marks the earliest occurrence of Planck's constant in physics. The method of derivation used here is due to Bose (1924).

It useful to define the variable

$$z = \frac{\hbar\omega}{kT}$$

as the ratio of $\hbar\omega$, the quantum unit of energy, to kT, the thermal unit of energy. In terms of z,

$$I(\omega)\,d\omega = \frac{\hbar}{\pi^2 c^3} \left(\frac{kT}{\hbar}\right)^4 \frac{z^3\,dz}{e^z - 1}.$$

The function $z^3/(e^z - 1)$ is plotted on the side of this page. It goes to zero as z^2 when $z \to 0$, and decays as $z^3 e^{-z}$ when $z \to \infty$. The function possesses a single maximum at $z \simeq 2.8214$. This immediately tells us that $I(\omega)$, the radiation's intensity, is maximum at $\omega = \omega_{\text{max}}$, where

$$\omega_{\text{max}} = (2.8214)\frac{kT}{\hbar}.$$

This remarkable scaling of the frequency at maximum with temperature is known as *Wien's displacement law*. It provides an efficient way of measuring the black-body temperature directly from its spectrum.

The energy density is now given by

$$u = \frac{\hbar}{\pi^2 c^3}\left(\frac{kT}{\hbar}\right)^4 \int_0^\infty \frac{z^3\,dz}{e^z - 1}.$$

This integral is one of the Bose-Einstein integrals encountered in Sec. F.11. It is equal to $\Gamma(4)\zeta(4) = 6\zeta(4)$, where $\zeta(x)$ is the Riemann zeta function. It turns out that $\zeta(4) = \pi^4/90$, so that u can be expressed in the simple form

$$\boxed{u = \frac{\pi^2}{15}\frac{(kT)^4}{(\hbar c)^3}}. \tag{7.6}$$

This remarkably simple result is known as the *Stephan-Boltzmann law*. It provides another way of measuring the black-body temperature.

7.5   Other thermodynamic quantities

We now calculate the pressure of a photon gas. We recall from Sec. F.1 that for a gas of massive particles, whether they are bosons or fermions, the pressure is always given by $P = 2E/3V$. When deriving this result, we saw that the factor of $2/3$ originated from the fact that for massive particles, the energy eigenvalues vary with volume as $\varepsilon_r \propto V^{-2/3}$. On the other hand, we recall from Sec. 2 that in the case of photons, the energy eigenvalues behave as $\varepsilon_r \propto V^{-1/3}$. We might therefore be tempted to guess that the pressure of a photon gas must be given by

$$\boxed{P = \frac{1}{3}u}, \tag{7.7}$$

where $u = E/V$ is the energy density. That this is in fact true follows from the same calculation as that presented in part f of Sec. F.1.

The entropy density of a photon gas, $s = S/V$, can be calculated by first computing $\ln \mathcal{Z}$ using Eqs. (6.12) and (7.1), and then substituting the result into the
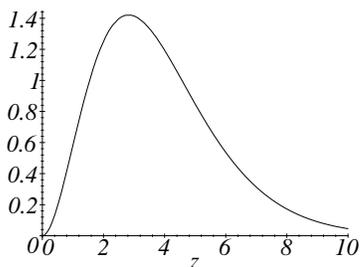
relation $S = k \ln \mathcal{Z} + E/T$. This derivation is the topic of a homework problem. Here we proceed differently, and deal directly with the first law of thermodynamics,

$$dE = T\,dS - P\,dV,$$

in which we substitute the relations $E = u(T)\,V$, $P = u(T)/3$, and $S = s(T)\,V$. Notice that we do not yet use the fact that $u(T)$ is given by Eq. (7.6). After some algebra, we see that the substitutions give

$$(4u/3 - Ts)\,dV + V(u' - Ts')\,dT = 0,$$

where a prime denotes differentiation with respect to T. Now, because V and T can be varied independently, we must have that each coefficient vanishes separately. This then tells us that

$$s = \frac{4u}{3T} = \frac{4\pi^2\,k}{45}\left(\frac{kT}{\hbar c}\right)^3, \tag{7.8}$$

which is the desired expression for the entropy density. Notice that the equation $u' = Ts'$, together with $s = 4u/3T$, confirm the fact that $u \propto T^4$.

Finally, the mean number density of photons, $n = N/V$, is computed from the relation $N = \sum_r \bar{n}_r$ by using all of the standard tricks. This also will be carried out in a homework problem. You should find that the number density is given by

$$n = \frac{2\zeta(3)}{\pi^2}\left(\frac{kT}{\hbar c}\right)^3, \tag{7.9}$$

where $\zeta(x)$ is our old friend the Riemann zeta function.

7.6   The cosmic microwave background radiation

It is a truly remarkable fact that our universe is filled with black-body radiation at a temperature of 2.7 K. This radiation, discovered by Penzias and Wilson in 1965, is known as the cosmic microwave background radiation (CMBR). A discussion of this phenomenon will conclude this section.

7.6.1   Recombination in the early universe

There was a time ($t < 10^5$ yr) in the history of the universe when atoms had not yet formed. The universe was then composed of light atomic nuclei, free electrons, neutrinos, and photons. By scattering them, the free electrons prevented the photons from propagating freely; the universe was opaque.

As the universe expanded, it cooled off, and eventually the electrons started to combine with protons to form hydrogen (and other light) atoms. The temperature of the universe at *recombination* can easily be estimated: kT must be somewhat smaller than the ionization energy of hydrogen, which is 13.6 eV. After conversion, this gives

$$T \sim 10^4 \text{ K.}$$

(A more precise calculation gives $T = 3575$ K.) Because of the disappearance of the free electrons, the photons could propagate freely after recombination; the universe became transparent.

The photons stayed in thermal equilibrium as the universe expanded to its present-day size. All the while, their temperature kept on decreasing. Today, this radiation (the CMBR) is measured to have a *perfect* black-body spectrum at a temperature

$$T_0 = 2.7 \text{ K.}$$

7.6.2   Dipole anisotropy

To a good approximation, the CMBR appears to be perfectly *isotropic*: the temperature is the same in all directions. (This poses quite a puzzle in cosmology, because it is not at all obvious that the entire universe had time to come to thermal equilibrium by the time recombination occurred. The currently favoured solution to this puzzle is known as inflation.) However, precise measurements reveal a slight variation with direction. This anisotropy is called the *dipole anisotropy*. Measurements show that the CMBR's temperature is more accurately described by

$$T(\theta) = T_0 + \Delta T \, \cos\theta,$$

where θ is the angle with respect to the direction at which the temperature is maximum. The amplitude of the temperature variation is $\Delta T = 3.2$ mK. So the CMBR appears to be hottest in one direction, and coolest in the opposite direction. Does this mean that there exists a preferred direction in the universe?

7.6.3   Black-body radiation as seen by a moving observer

Before we answer this question, let us ask how black-body radiation would appear to a *moving* observer. Suppose that a photon emitted by a black body at a temperature T_0 has a frequency ω_0 as measured by an observer at rest with respect to the body. This photon is now measured by an observer which is moving with respect to the body. If the observer moves *directly toward* the body with a velocity v, then the measured frequency ω will be slightly larger (shifted toward the blue) than ω_0: $\omega = \omega_0(1 + v/c)$. (This is just the good old Doppler effect.) If, on the other hand, the observer moves *directly away from* the body, then the measured frequency ω will be slightly smaller (shifted toward the red) than ω_0: $\omega = \omega_0(1 - v/c)$. More generally, if the observer moves at an angle θ with respect to the black body, then the measured frequency is

$$\omega = \omega_0\left(1 + \frac{v}{c}\,\cos\theta\right).$$

This dependence on θ looks suspiciously similar to the one encountered before.

The spectrum of the black-body radiation, as measured by the observer *at rest*, is given by the Planck function, which depends on the quantity

$$\frac{\hbar\omega_0}{kT_0}.$$

When expressed in terms of ω, the frequency measured by the *moving* observer, this quantity becomes

$$\frac{\hbar\omega}{kT_0(1 + \frac{v}{c}\,\cos\theta)}.$$

This shows that the spectrum, as measured by the moving observer, is also Planckian, but with a temperature

$$T(\theta) = T_0\left(1 + \frac{v}{c}\,\cos\theta\right).$$

This temperature, as that of the CMBR, depends on the angle of approach: A black body appears cooler when the observer is moving away from it, and hotter when the observer is moving toward it.

7.6.4   Proper motion of the galaxy

So, does the universe possess a preferred direction? The answer is no. The $\cos\theta$ variation in temperature is due to the Earth's relative motion with respect to the rest frame of the CMBR (the rest frame of the universe!). The Earth's velocity v_E is easily obtained from the relation $\Delta T/T_0 = v_E/c$, which gives

$$v_E = 355 \text{ km/s}.$$

Part of this motion is due to the rotation of the sun around the galactic centre. But this does not account for the whole effect: the remaining part must be attributed to the motion of the galaxy as a whole:

$$v_G = 500 \text{ km/s}.$$

To explain this proper motion of our galaxy is still a major challenge of cosmology.

7.6.5   Quadrupole anisotropy

Once the dipole anisotropy has been subtracted out, the CMBR appears almost perfectly isotropic. There remains, however, a slight anisotropy (the *quadrupole anisotropy*) of amplitude

$$\frac{\Delta T}{T_0} \sim 10^{-5}.$$

This remaining anisotropy gives an important clue as to how galaxies might have formed in the early universe. This is the topic of a lot current work in cosmology.

7.7   Problems

1. Calculate the following quantities for a photon gas in thermal equilibrium:

   a) The entropy density $s = S/V$. [You must calculate s directly, without referring to the thermodynamic derivation presented in the notes.]

   b) The mean number density $n = N/V$.

2. A photon gas at a temperature T_A fills a thermally insulated cavity of volume V_A. If the cavity is expanded reversibly to a new volume V_B, what is T_B, the final temperature?